119 p.

Report No. SSD 3166R) OTS: DCOCS-3

Prepared for:

NATIONAL AERONAUTICS AND
SPACE ADMINISTRATION
MANNED SPACECRAFT CENTER
HOUSTON 1, TEXAS

STUDY ON OPTICAL COMMUNICATIONS FROM DEEP SPACE

INTERIM PROGRESS REPORT

1 February 1963 through 27 March 1963

HUGHES

AEROSPACE GROUP
SPACE SYSTEMS DIVISION
HUGHES AIRCRAFT COMPANY
CULVER CITY, CALIFORNIA

(NASA Contract No. NAS 9-879)
P. R. No. 3-440-20351
WAS 10-81-02-36

auth of V

Report No. SSD 3166R DCOCS-3

Prepared for:

NATIONAL AERONAUTICS AND
SPACE ADMINISTRATION
MANNED SPACECRAFT CENTER
HOUSTON 1, TEXAS

STUDY ON OPTICAL COMMUNICATIONS FROM DEEP SPACE

INTERIM PROGRESS REPORT

1 February 1963 through 27 March 1963

Contract No. NAS 9-879
P. R. No. 3-440-20351
WAS 10-81-02-36

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

INTRODUCTION

16727

This Interim Progress Report presents a discussion of the work performed on Contract NAS 9-879 during the period from 1 February 1963 through 27 March 1963.

The modulation and de-modulation task is fully reported in this volume. While the principal emphasis in that work has been placed on an examination of wide bandwidth systems with the goal of providing real time TV capability, the feasibility of utilizing such capability at these extreme ranges is strongly dependent on the state-of-art of other system parameters. As a result, techniques for providing narrower bandwidths are also discussed.

A treatment of atmospheric attenuation studies is then given, using knowledge gained from the Hughes Aircraft Company Malibu-Baldwin Hills link. This is followed by a discussion of the extremely important pointing problem, the result of which is seen to have considerable impact on the design bandwidth. Frequency selection is also treated in a preliminary way.

The report concludes with a treatment of the system implementation problem and discusses the advantage and disadvantages of the various communications links. Some preliminary conclusions for system selection can be drawn, providing a basis for the system design which is the principal subject of the next report period.

AUTHOR

Kenneth L. Brinkman

KENNETH L. BRINKMAN [1963] 119 pgs.
Project Manager
NASA Deep Space Optical
Communications Study

1.0 TYPES OF MODULATION

Many techniques are available to permit the modulation of light beams at low frequencies. Most of them have been known and put to practical use long before the advent of the laser.

The purpose of the following section is to survey some of the various modulation methods applicable to the different types of lasers available and to make specific recommendations for their use.

The method of modulation will depend strongly on the type of laser used. For a semiconductor laser, for instance, pump power modulation to video bandwidths is a straight forward and simple problem. The nature of the problem is radically different in a gas laser or ruby laser. Similarly the success of a frequency modulation system will depend on the frequency stability of the laser and the frequency spacing between longitudinal modes. Spiking on the other hand may hinder intensity modulation.

1.1 INTENSITY MODULATION

There exist various methods to modulate the intensity of a laser beam. The two most promising from a communication standpoint will be considered here. These are the electro-optical effect, in which use is made of the change of the optical properties of certain materials under the strain of an electric field; and pump modulation in which the laser pump mechanism is altered.

1.1.1 Electro-optical Effect

Certain liquids and crystals can change plane polarized light into elliptical polarization when an electric field is applied to them either transverse or longitudinally to the light propagation. If the optical medium is a liquid the phenomenon is known as the Kerr effect, if a crystal it is the Pockels effect.

The change of polarization is basically caused by the dependence of the index of refraction, and thus of the propagation constant along certain optical axes of the optical media on the applied electric field. A variation of the propagation constant along the optical x-axis, for instance produces a change in phase with respect to a beam polarized along the y-axis, thus inducing different degrees of elliptical polarization.

In liquids the changes of index are based upon a statistical preferential alignment of dipolar molecules in an electric field. The upper frequency limit is determined by the relaxation effects and losses occurring when the inverse frequency becomes comparable to the characteristic correlation time for Brownian motion. In the case of nitrobenzene this time is about 50 psec. In most solids

the characteristic time is much longer or does not exist. However certain crystals such as hydrogen phosphates allow displacements in their atomic structure which account for the Pockels effect. The upper frequency limits of such crystals has not yet been fixed. A practical limitation of the Kerr effect as a light modulator is that the induced optical retardation is proportional to the square of the applied electric field, whereas in a Pockels cell the effect is linear. This results from the fact that in liquids all of the molecules are free to rotate, so one has to average over all orientations of the molecules and the linear terms drop out. In solids the linear terms do not average out. Only under strong biasing electric fields will a Kerr cell's phase retardation behave linearly. In the best cases the electro-optical constant of a Kerr cell is similar to that of a Pockels cell. This coupled to the fact that liquid cells require higher purity to conserve transmission properties and add complexity to light modulating structure seem to make Pockels cells more useful as light modulators than Kerr cells.

The hydrogen phosphate Pockels cells have fairly good light transmission properties in the spectral region from 0.3 to 1.1 microns. The percent transmission in this region averages higher than 80% for 1/4 inch crystal thickness. Cuprous chloride or cuprous bromide which have also been used as electro-optical modulators reportedly transmit light out to wavelengths of 15 microns.

Not all of the light transmitted through the hydrogen phosphate crystals is modulated. A considerable fraction of the light "leaks" through the system because of the anisotropic nature of the dielectric constant and slight dislocation of the optical axis. The light leakage is a function of the solid angle of the beam. When a Pockels cell is placed between crossed polarizers, the ratio between maximum and minimum transmission is defined as the contrast ratio. For a contrast ratio of 100/1 in a hydrogen phosphate crystal of 1/16 inch thickness and transmitting light at 0.5 microns, the maximum acceptable angle of divergence of the beam is 2.5 degrees. This cone angle is inversely proportional to the square root of thickness times wavelength. For the highly collimated laser beams to be modulated light leakage can be held to a minimum. The cuprous chloride and zinc sulphide Pockels effect crystals are cubic and, in the absence of an electric field, isotropic. This means that light leakage is low and thus, for a given beam divergence, much longer crystals may be used. This, as will be shown when studying traveling wave modulators, (section 2.1.2), is a useful property where long interactions between light beams and electric fields are required.

Figure 1-1 shows a typical arrangement of a Pockels cell intensity modulator. The quarter wave plate biases the incoming linearly polarized light to circular polarization. The end polarizers are crossed. Under such conditions it can be shown that

$$\frac{I}{I_o} = \frac{1}{2} + \sin\left(\frac{\pi}{\lambda} n^3 r V \sin \omega t\right) \qquad (1)$$

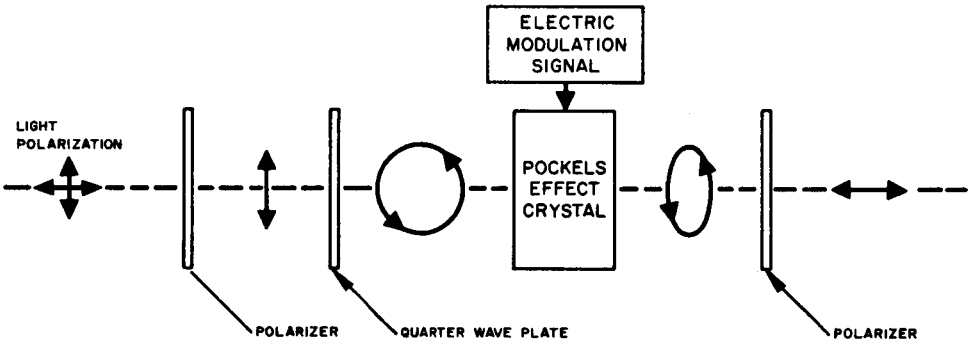

Figure 1-1. Pockels Cells Light Intensity Modulator

where I_o is the incoming light beam intensity, I is the modulated intensity, λ is the light wavelength, n is index of refraction of the ordinary beam in the Pockels cell, r is the electro-optical constant, V is the modulating voltage and ω is the frequency of modulation.

Using a Bessel function expansion equation (1) becomes

$$\frac{I}{I_o} = \frac{1}{2} + 2\left[J_1(a)\sin\omega t + J_3(a)\sin 3\omega t + J_5(a)\sin 5\omega t - - -\right]$$

where

$$a = \frac{\pi}{\lambda} n^3 r V$$

The total percent distortion is defined as

$$D = \frac{100}{J_1(a)}\sqrt{\left[J_3(a)\right]^2 + \left[J_5(a)\right]^2 + - - -}$$

Since there is no second harmonic the distortion will always be small for small values of a, Table 1-I shows the percent distortion and modulation for different values of a. In this case the percent modulation is defined as

$$m = \frac{I_p/I_o - 0.5}{0.5}$$

where I_p is the peak intensity. At 100% modulation the distortion is still less than 1%.

If the quarter plate in Figure 1-I is removed equation (1) becomes

$$\frac{I}{I_o} = \sin^2\left[\frac{\pi}{\lambda} n^3 r V \sin\omega t\right]$$

which again may be expanded to take on the following form

$$\frac{I}{I_o} = 4\left\{\left[J_1(a)\right]^2 + \left[J_3(a)\right]^2 + - - - + \left[J_1(a)\,J_3(a) - \frac{\left[J_1(a)\right]^2}{2}\right]\cos 2\omega t - J_1(a)\,J_3(a)\cos 4\omega t - \frac{J_3(a)}{2}\cos 6\omega t + - - -\right\}$$

in which case the fundamental modulation frequency disappears.

Pockels cell modulators may be operated at room temperature. However the crystal's optical properties are susceptible to temperature changes and care should be taken to maintaining it at a constant value. This causes, in many cases, heat dissipation problems since crystals are usually small and modulation power requirements are relatively high. By subjecting the Pockels cell to low ambient temperature an improvement in performance results. When temperature is lowered to the Curie point of the crystal the electro-optical properties are enhanced. Typical results indicate a 10 to 1 reduction in the half-wave retardation voltage of hydrogen phosphate crystals.

In most applications of Pockels crystals to light modulators the end surfaces must be polished to optical flatness and with the optical axis oriented to an accuracy of better than $\pm$ 10 seconds. Achieving these accuracies is often complicated by the fragile nature of the crystals.

An additional problem is the reflection loss at the end surfaces. To minimize this it is necessary to optically match the crystal with a substance of similar dielectric constant.

1.1.2 Pump Power Modulation

The electro-optical modulation considered in 1.1.1 took place outside of the optical cavity. The means by which the beam to be modulated was produced was of secondary interest. The modulation to be considered in this section takes place inside the cavity and the characteristics of the cavity will determine some of its limits. Therefore a short consideration of the laser optical cavity will be useful.

In the optical cavity, a resonance frequency f_r will occur when the cavity is an integral number of half wavelengths long.

$$f_r = \frac{cM}{2nL} \tag{1}$$

where c is the speed of light in a vacuum, M is an integer, n the index of the medium inside the cavity and L is the mirror spacing. M may be interpreted as the mirror spacing measured in half wavelengths. Optical spacings, nL, of 100 cm and 10 cm result in resonances spacing of 150 mc and 1.5 kmc respectively. The width of the resonance between half power points for a passive cavity, is given to a good approximation by*

$$\Delta f_c = \frac{(1 - R_p) f_r}{M^{1/2}} \tag{2}$$

* Vern N. Smiley, "An Active Interference Filter as an Optical Maser Amplifier, Proc. IRE, Jan. 1963.

W. R. Bennet, Jr., "Gaseous Optical Masers," Applied Optics, Supplement on Optical Masers, 1962.

where R_p is the Fresnel reflection coefficient. Combining (1) and (2):

$$\Delta f_c = \frac{(1 - R_p)c}{2\pi nL} \tag{3}$$

for nL = 100 cm and $R_p = 0.99$, the bandwidth is approximately 0.5 mc.

When a laser material is placed inside the cavity, with a single pass gain of G, (3) must be replaced by

$$F_c = \frac{(1 - GR_p)c}{2\ nL(GR_p)^{1/2}} \tag{4}$$

Minimum oscillation condition is given by $GR_p = 1$. Actually, oscillation will begin at GR_p slightly less than one due to spontaneous emission.

The maximum transmission, T_o, neglecting absorption losses in the end mirrors, may be written

$$T_o = \frac{(1-R_p)^2 G}{(1-R_p G)^2} \tag{5}$$

The gain of the laser material is controlled by the intensity of the pumping source. The characteristics of the laser, then, as it starts to oscillate are

a) The emission frequency bandwidth narrows.

b) The intensity of radiation within this bandwidth increases.

The inherent bandwidth of the laser has been determined experimentally* to be less than 2 cps. However, due to thermal and mechanical variations in the cavity length, the output frequency of the laser oscillator varies over several tens of kc, even with electronic tuning and feedback correction.

From (5), the power output is proportional to the gain of the laser system, which in turn is dependent on the pumping power. It is this mechanism which makes possible amplitude modulation of the pumping power. A typical curve of power out vs. pumping power is shown in Figure 1-2.

---

*A. Javan, et al, "Frequency Characteristics of a Continuous-wave Heue Optical Maser," Journ. Am. Opt. Soc. p 46, Jan. 1962.

At some threshold value of pumping power P_1, the tube starts to emit a coherent beam which increases thereafter with increases in applied power. At some point P_2, the power-out versus power-in curve will level off or start to decrease depending on the composition of the individual laser.

Of primary interest, when pump power modulation is being considered, is the shape of the transfer function curve in Fig. 1-2. The point at which the curve drops to zero, P_1, represents the minimum instantaneous value of pump power which can be used before the laser extinguishes. The point at which the curve flattens out or starts to decrease with increasing pumping power, P_2, represents the beginning of nonlinearity and distortion. The region between P_1 and P_2 then represents the range over which the pumping power may be varied with a minimum nonlinearity.

The percent distortion in the modulated beam will depend on the deviation of the curve from a straight line between P_1 and P_2, which in turn depends on the particular laser composition.

If the average pumping power, P_o, is

$$P_o = \frac{P_1 + P_2}{2}$$

then

$$\% \text{ modulation} = \frac{P_o - P_{min}}{P_o} \times 100 = K \frac{P'_o - P'_{min}}{P'_o} \times 100$$

if the transfer curve is linear. If $P_{min} = P_1$, this is the maximum pump modulation level which may be used without extinguishing the laser action. The % modulation is not 100% unless $P_1 = 0$, which is not the case. For most lasers the modulation level is limited to considerably less than 100%.

The efficiency of the modulation process is a second consideration. Unlike some modulation processes which require very little power by utilizing the gain of the cavity, the pump modulator must supply power on the order of magnitude of the pump source itself and the resulting modulation efficiency is approximately equivalent to the pumping efficiency.

Since part of the discussion of pump power modulation is peculiar to the type of laser under consideration, these will be discussed separately.

a - Semiconductor Laser

The semiconductor laser, such as the Ga As diode, is pumped by a low voltage, high current d-c power supply. Both pulsed and CW operation of these devices has been reported. The pump modulation of this type of laser provides

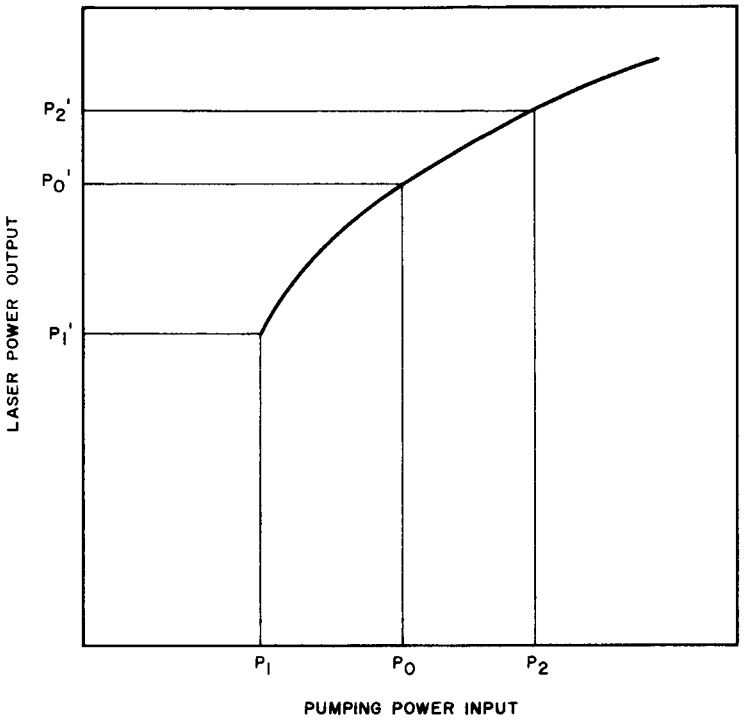

Figure 1-2. Typical Relationship between Laser Power Output and Pumping Power Input

the most promising means of control. Upper frequency limitations are not imposed by pumping frequency, as in r-f excited lasers, but by the physical characteristics of the device itself. Modulation bandwidth is expected to be relatively high. The incoherent emission of a Ga As diode pumped below threshold has already been used to transmit a video picture.

b - Gas Laser

The gas laser becomes more involved than the semiconductor laser to pump modulate. The gas laser generally operates CW, although pulsed operation is also possible, and is generally excited by an r-f source, although, again, d -c tubes are under development.

The r-f pumping frequency becomes a kind of subcarrier frequency which must be amplitude modulated and imposes an upper frequency limitation on the modulating signal.

Several mechanisms are responsible for the population inversion in the gas laser, one of which is a collision between an excited He atom and the Ne atom. The diffusion time for the gas atoms sets a second and more stringent upper frequency limitation on the modulation rates. This plus other effects presents an inherent limitation to video frequencies*.

Although other effects may overshadow it, the bandwidth of the optical cavity will also present a frequency limitation on the modulation. The half power bandwidth of a typical gas laser with a 1 meter cavity may easily be 1 mc or less. If both sidebands are to be kept within this frequency, the modulating frequency must be half, or less, of the cavity bandwidth. From (2) the maximum modulating frequency, F_{max}, can be written

$$F_c = \frac{(1 - R_p)\, c}{2\pi nL\, R_p^{1/2}} .$$

For nL = 100 cm and R_p = 0.99, the maximum modulating frequency is 0.25 mc.

c - Solid State Laser

The solid state laser, such as the ruby, generally operate in a pulsed mode. However, continuous operation is possible if the excess heating power can be dissipated. Theoretical considerations lead to the conclusion that the maximum information transmission rate is approximately equal to

*P. S. Pershan and H. Bloemberger, "Microwave Modulation of Light," Advances in Quantum Electronics, Columbia University Press, 1961, p. 187-199.

$$R \leq (\Delta f_c) \ln \left[\frac{S}{N} + 1 \right]$$

or if amplitude modulation is disregarded

$$R \leq \Delta f_c$$

For a 10 cm cavity at $\lambda = 1\mu$, 2% loss,

$$f_c = 10 \text{ mc}$$

However, solid state lasers have only been pulsed at rates of at most, a few pulses per minute. Obviously this is much too slow to be useful in information transfer. Most of the time is absorbed in changing the flash lamp used to excite the laser crystal. With "hair-trigger" operation in which the main flash lamp excites the laser to just below threshold and a supplementing excitation causes lasering, pulse repetition rates of the 1 kc have been maintained for the duration of the flash lamp output (1 sec.). Even at this frequency, the pulses cannot be maintained continuously. This requires that the information to be transmitted, which does not arrive in pulses, must be stored between transmission bursts. If we consider a 1 megabit/sec signal rate, a pulse repetition frequency of 1/sec and a duty cycle of 1 m sec, then each pulse must contain 1 million bits of information which must be read onto the pulse in 1 msec. This requires a storage buffer of 1 megabit capacity and a readout speed of 1000 megabits/sec. This is far beyond present capabilities.

The main attribute of the pulsed ruby laser, is their large instantaneous power output, which may be in the hundreds of kilowatts. This is compared to a few milliwatts for the c-w gas laser or a few hundred milliwatts for the pulsed gas laser.

1.1.3 Pressure Modulation

This technique involves the establishment of standing wave patterns in some transparent crystals, such as glass or quartz. Due to photoelastic effects, the alternate tension and compression areas between modes change the polarization of the light traveling through it. If this plate is used in conjunction with polarized analyzer plates, this rotation shows up as amplitude modulation of the light beam. Both tension and compression areas will produce a light output, causing doubling of the modulating frequency. This frequency is used as a subcarrier, on which information is impressed, and may be either amplitude or frequency modulated. Frequency bandwidth is limited to about 100 kc in a 6 inch block.

The advantage of this method is that relatively cheap glass or quartz may be used as the modulator, requiring only 0.5 watt/in^3 for glass and 0.01 watt/in^3 for quartz. Such modulators have been built by Decker Aviation Co. using carrier frequencies up to 10 mc.

1.2 POLARIZATION MODULATION

Polarization modulation may be obtained by using the Pockels effect in section 1.1.1. The light beam emerging from the Pockels crystal in Figure 1-1 is elliptically polarized. By inserting an analyzer 50% of the transmitted power is lost since only one component of the elliptical polarization is used. This loss may be eliminated by transmitting the light beam as it emerges from the Pockels crystal. At the receiving end a device which could distinguish between the elliptical polarization components would then be necessary. Such a device is a Wollaston Prism. A Wollaston Prism is composed of two prisms. Light enters normal to the surface of the first prism and travels perpendicular to its optic axis until it strikes the second prism where double refraction takes place. The two emerging beams are polarized at 90^{o} with respect to one another and thus provide the two components of the elliptical polarization when properly oriented.

Polarization modulation has several advantages over intensity modulation. Besides transmitting twice as much power there is a significant improvement in S/N at the receiver in the presence of background noise.

All of the other properties pertaining to the Pockels cell mentioned in section 1.1.1, are true also in the case of polarization modulation.

1.3 FREQUENCY & PHASE MODULATION

1.3.1 Electro-Optical Effect

Again the electro-optical effect discussed in section 1.1.1 may be used for either frequency or phase modulation of light. The basic limitations and shortcomings of the Pockels cell described above are still the case. The principles for frequency modulation and phase modulation will be described below using the Pockels effect. Methods for detecting these types of modulation will be described in section 4.

a - Frequency Modulation

It can be shown that when a Pockels crystal is subjected to a rotating electrical field circularly polarized light propagating through the crystal will experience a shift in frequency equal to that of the rotating electric field*.

---

*C. F. Buhrer et al. "Optical Frequency Shifting by Electro-Optic Effect," Applied Phys. Letters, Vol. 1, N. 2, P. 46.

Consider a left-hand circularly polarized beam (circular polarization is obtained by having light travel through a polarizer and quarter wave plate such that the plane of polarization is at $\pm 45^{\circ}$ with the axes of the quarter wave plate; the sense of the rotation is determined by the sign of the 45° angle). The electric vector of the incident light is

$$A_x = A \cos \omega t \quad , \quad A_y = A \sin \omega t \tag{1}$$

where x and y are the coordinate axes of the crystal. The light is assumed to be traveling along the Z, or optic, axis. On application of a left hand rotating electric voltage in the plane perpendicular to the z axis, with components

$$V_x = V_m \cos \omega_m t, \qquad V_y = V_m \sin \omega_m t \tag{2}$$

the light emerging from the crystal will take on the form

$$\begin{aligned} A_x &= A \cos \frac{\delta}{2} \cos \omega t - A \sin \frac{\delta}{2} \sin \left[(\omega + \omega_m)t + \phi \right] \\ A_y &= A \cos \frac{\delta}{2} \sin \omega t - A \sin \frac{\delta}{2} \cos \left[(\omega + \omega_m)t + \phi \right] \end{aligned} \tag{3}$$

where

$$\delta \cong \frac{2\pi n^3 r V}{\lambda} \tag{4}$$

is the retardation and ϕ is a fixed angle independent of the field. The first terms in A_x and A_y correspond to a left-circularly polarized vibration at the original frequency with an amplitude $A \cos \frac{\delta}{2}$. The second terms in A_x and A_y correspond to a right-circularly polarized vibration at the frequency $\omega + \omega_m$ with amplitude $A \sin \frac{\delta}{2}$. If δ can be made qual to π, all of the light is transmitted with a right polarization shifted to the frequency $\omega + \omega_m$. Otherwise it would be necessary to use a right hand polarizer at the output or receiving end, and thus introduce losses into the system. It should be noted that if the direction of either polarization rotation (light beam or electric field) were reversed, the frequency $\omega - \omega_m$ would be obtained.

Consider now that the left hand polarization has been eliminated and the

rotating voltage is frequency modulated, i.e.

$$\omega_m(t) = \omega_m t + M \sin \omega_f t$$

substituting in (3)

$$A_x = B \sin (\omega' t + M \sin \omega_f t)$$
$$A_y = B \cos (\omega' t + M \sin \omega_f t) \qquad (5)$$

where $B = -A \sin \frac{\delta}{2}$, $\omega' = \omega_m + \omega_f$ and $\emptyset$ can be assumed to be zero. Going through the conventional Bessel function expansion used to study the frequency spectrum of an FM wave (5) may be expressed as

$$A_x = J_o(M)\, B \sin \omega' t + J_1(M)\, B \left[\sin (\omega' + \omega_f) t - \sin (\omega' - \omega_f) t\right] + J_2(M)\, B \left[\sin (\omega' + 2\omega_f) t + \sin (\omega' - 2\omega_f) t\right] + - - - -$$

$$A_y = J_o(M)\, B \cos \omega' t + J_1(M)\, B \left[\cos (\omega' - \omega_f) t - \cos (\omega' + \omega_f) t\right] + J_2(M)\, B \left[\cos (\omega' - 2\omega_f) t - \cos (\omega' + 2\omega_f) t\right] + - - -$$

The bandwidth of the modulator therefore depends on the magnitude of the different order Bessel function coefficients of the sidebands which in turn are a function of their argument M. The maximum allowable modulator bandwidth will be dependent on the type of laser used and its longitudinal mode of operation. Since all of the possible frequency components of the laser beam will be equally modulated there is the possibility of interference between overlapping sidebands.

b - Phase Modulation

A similar method of obtaining phase modulation of a laser beam using the electro-optical effect will now be described. This modulator impresses phase modulation on the light by modulating the velocity of propagation through the electro-optic material. It can be shown that if light is plane polarized in the direction of the x crystal axis and traveling along its y axis with an electric field applied across the z axis the phase retardation will be*

* C. J. Peter, "Gigacycle Bandwidth Coherein Light Traveling Wave Phase Modulator"

$$\delta = \frac{\pi}{\lambda} n^4 r V \tag{7}$$

where the parameters are as defined in Section 1.1.1.

The phase retardation δ is seen to be directly proportional to the applied voltage. The light traveling through the crystal will maintain its plane polarization along the x axis of the crystal. The phase modulated light will introduce bandwidth limitation similar to those described for frequency modulation. If the incoming electric light vector is of the form

$$A_x = A \sin \omega t$$

the energy beam will be

$$A_x = A \sin (\omega t + \delta) = A \sin (\omega t + N \sin \omega_f t) \tag{8}$$

where in this case the modulating voltage is

$$V = V_m \sin \omega_f t$$

and

$$N = \frac{\pi}{\lambda} n^4 r V_m \tag{9}$$

when expanded into a bessel function series (8) becomes

$$A_x = J_o (N) A \sin \omega t + J_1 (N) A \left[\sin (\omega + \omega_f) t - \sin (\omega - \omega_f) t \right] + J_2 (N) A \left[\sin (\omega + 2\omega_f) t + \sin (\omega - 2\omega_f) t \right] + - - - \tag{10}$$

In the case studied above of frequency modulation the modulating electric field had to be frequency modulated. For phase modulation the applied field must be amplitude modulated. The magnitude of N will in most cases be less than one. A typical Pockels crystal such as KDP for instance has an electro-optical coefficient $\gamma = 11 \times 10^{-7} \frac{cm}{Kv}$ and an index of refraction $n = 1.468$; thus at a light wavelength $\lambda = 0.5$ microns the required modulating voltage is from (9)

$$V_m = \frac{N \lambda}{\pi n^4 \gamma} = \frac{0.5 \times 10^{-4} N}{\pi \times 1.468^4 \times 11 \times 10^{-7}} = 3.1 (N) \quad KV$$

More than 3 Kv are necessary to make N = 1 in this case. For the Bessel function coefficients are negligible beyond the first order and therefore second harmonic sideband interference between the different laser longitudinal modes is very small. Thus for multimode laser operation it may be concluded that negligible distortion results for modulating frequency $w_f < \frac{\Delta F}{2}$ where ΔF is the frequency separation between longitudinal laser mode.

This result should be contrasted with the bandwidth limitations of frequency modulation. In equation 6 the argument of the Bessel functions is M, the frequency deviation ratio of the frequency modulated electric field. If M is much larger than 1 bandwidth is considerably restricted as compared to phase modulation. In any event, in both phase and frequency light modulation the laser beam must be maintained fairly stable, a difficult problem with most types of lasers.

1.3.2 Zeeman Effect

The application of a magnetic field, H, results in the splitting of the normal laser transition line into two transition lines, separated from the normal b y plus or minus a frequency which is proportional to the magnetic field strength, and circularly polarized in opposite directions. This effect, which has been treated as a laboratory curiosity, might be used to modulate a laser beam in either polarization angle, frequency or amplitude. The Zeeman effect would be equally applicable to any type of laser, but for simplicity, the He Ne laser, 1 meter long, with plane parallel mirrors will be considered.

When the laser is weakly excited, so that only one longitudinal mode is excited, and in the presence of a static longitudinal magnetic field, its emission is a doublet of right and left circular polarization. The doublet may belong to adjacent cavity modes ($\sim$150 mc apart) or to the same mode. The superposition of the doublet results in a linearly polarized wave oscillating at the average of the doublet frequencies and with its plane of polarization rotating at half the difference frequency. When viewed through a linear polarizer, the wave appears amplitude modulated at the difference frequency, regardless of the orientation of the polarizing plate.

Since the difference frequency is linearly dependent on the magnetic field, the output wave is of the form

$$e = E \sin \omega_c t$$

where $\omega_c = KH$, K = constant, H the magnetic field.

If the field, H, is modulated around zero the output is amplitude modulated at twice the modulating frequency due to the zero H field. However, if a d-c bias field is applied and the H field modulated as sin ω_m t:

$$\omega_c t = \omega_o t - \Delta\omega \sin \omega_m t$$

$$e = E \sin (\omega_o t - m \sin \omega_m t)$$

which is the equation for a frequency modulated wave. However, the carrier, ω_o, is not the optical frequency, but is determined by the applied d-c field.

The maximum frequency deviation is limited by the bandwidth of the optical cavity which can be as low as 0.5 mc. In a typical He Ne laser, the frequency of rotation of the plane of polarization

$$\delta f/_N = h^{-1} \, gB \, \frac{\Delta f_c}{\Delta f_{Ne}} = 600 \text{ cps}/_{\text{oersted}}$$

where h = Planck's constant, g = splitting factor, B = Bohr magneton, Δf_c and Δf_{Ne} are the half power widths of the laser cavity and of the Ne transition line respectively, 600 cps is about the minimum frequency shift.

The Zeeman effect may thus be used to directly frequency modulate the laser beam if a circular polarization detector, such as a quarter wave plate and linear polarizer, is used to detect the individual frequency components of the Zeeman doublet. Frequency detection then must take place at optical frequencies or super heterodyne techniques must be used to convert to an IF frequency.

If the laser is strongly enough excited so that more than one longitudinal mode can build up at one time, amplitude modulation may be detected without a polarizer plate. The laser will than operate in as many as 3 modes at once. These modes contain circularly polarized waves at different frequencies, which, when superimposed, produce an amplitude modulated wave. The modulation will be $\Delta f = \frac{c}{2L}$, the spacing of resonant frequencies in the optical cavity. For a 1 meter cavity, Δf = 150 mc.

1.3.3 Magnetostrictive

When a Ferromagnetic material is placed in a magnetic field, its length changes in relation to its magnetostriction constant. This constant may be either positive or negative, that is the material may either expand or contract in the presence of a magnetic field. If the cavity containing the end mirrors were constructed of a ferromagnetic material, the magnetostrictive effect could be used to control the spacing and alignment of the mirrors. Various methods have been suggested for accomplishing this.*

* W. R. Bennet, et. al, "Magnetostrictively Tuned Optical Maser," Rev. Sci. Instr. 33:601-5, June, 1962.

For any resonant cavity, the resonant frequency may be represented by:

$$f = \frac{M c}{2L}$$

Where C is the speed of light, M is an integer and L is the reflector separation. For a given change in cavity length, ΔL, the change in resonant frequency, Δf, is

$$\Delta f = f_1 - f_2 = \frac{M_1 c}{2L_1} - \frac{M_1 c}{2(L_1+\Delta L)} = \frac{M_1 c}{2}\left(\frac{\Delta L}{L_1(L_1+\Delta L)}\right) \cong \frac{M c}{2L_1}\left(\frac{\Delta L}{L_1}\right) \tag{1}$$

or $$\Delta f = f_1 \left(\frac{\Delta L}{L_1}\right)$$

The change in frequency is directly proportioned to the change in cavity length.

The normal laser excitation within the cavity is much broader than the cavity bandwidth. Therefore it is possible for more than one frequency to oscillate at once. In the plane-parallel Fabry-Perot cavity, one meter long, these longitudinal modes are separated by

$$f = \frac{c}{2L} = 150 \text{ mc}$$

Evidently then, if this laser is to be frequency modulated, the frequency deviation must be less than 150 mc, in order to be able to suppress these unwanted components. On the other hand, thermal and mechanical effects in the mirror supports produce "noise" of the order of 100 kc so the frequency deviation must be greater than this to give good signal to noise ratio. For the same cavity and frequency considered above, the length change for a frequency deviation of 1 mc/sec is:

$$\frac{\Delta L}{L} = \frac{\Delta f}{f_1} = \frac{10^6}{3 \times 10^{14}} = \frac{1}{3 \times 10^8}$$

This much change is easily achievable in many magnetostrictive materials.

The magnetostrictive effect varies with the material, positive in some and negative in others. Some materials exhibit negative coefficients over a portion of their range and a positive coefficient over the rest. The sign of the coefficient is the same for either sense of magnetization, so that a DC bias current would have to be used to eliminate large non-linear distortion.

Consider iron as a magnetostrictive material. It's curve appears in Fig. 1-3. If this were the magnetostrictive material used to move the mirrors, and if a DC field of strength H_o were applied, there would be no change of cavity length in the absence of a modulating signal. An applied signal would cause the cavity length to be changed proportional to the applied field. This material gives a negative coefficient centered on zero change of length. Other materials exhibit a positive coefficient in this region and others start at zero but do not recross the axis. The material should be chosen for the particular application, bearing in mind that its susceptibility to mechanical and thermal variation will also affect the steady state frequency stability. For the small signal input required, distortion due to nonlinearity in the magnetostriction coefficient will be very low.

Magnetostrictive elements have been used as oscillator tuning elements, at frequencies generally below 70 kc. They are also used in delay lines at frequencies up to 4 mc. However, the weight of the mirror assembly which the rods must move will greatly effect their frequency response and frequencies in excess of 1 mc seem unlikely. Furthermore, the stability of the biasing field would greatly affect the noise on the laser beam, and mirror alignment must be maintained, while the beam is being modulated, to an accuracy of better than 6" or arc. Feedback methods might be used to stabilize the cavity and reduce low frequency noise on the laser beam.

Magnetostrictive techniques can also be used to amplitude modulate the laser by changing thealignment of the cavity mirrors. This, however, encounters the same internal frequency limitations which limit pump power modulation. For maximum modulation, mirror misalignment of about 6" of arc is necessary -- this is easily attainable.

1.4 COMPARISON OF MODULATION METHODS

Table 1-II shows a comparison of advantages and disadvantages between some of the different modulation methods discussed in this section. On the basis of this table it would appear that polarization modulation offers the most promising possibilities for solid state and gas lasers and pump power modulation for semiconductor lasers.

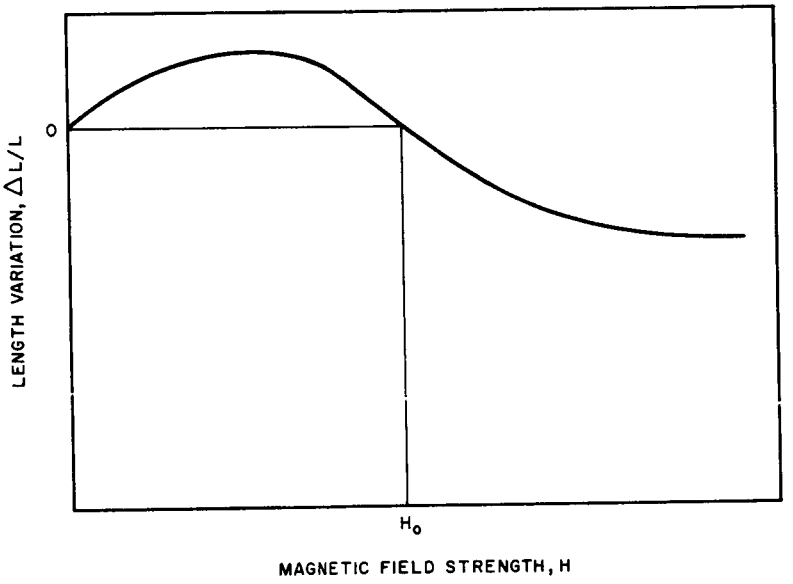

Figure 1-3. Magnetostriction of Iron

TABLE 1-II COMPARISON OF MODULATION METHODS

Type	Advantages	Disadvantages
Intensity using Pockels Effect	1. Low distortion for large modulation index (1% for 100% modulation). 2. Potential high modulation frequency (10 gc). 3. Transmission attenuation is low in crystal	1. Requires relatively high power (600 v, 0.5 amps) at video frequencies 2. Beam must be well collimated to avoid light leaks. 3. Crystals are temperature sensitive and easily damaged. 4. Spectral transmission for most crystals is limited to 1.1 microns. 5. Reflection loss may be high if crystal is not optically matched. 6. Transmission loss of 50% due to polarizer
Intensity using Pump Power	1. For semiconductor lasers video bandwidths are possible. 2. Method is simple and straightforward. 3. Modulation power requirements are of same order of magnitude as pump power.	1. Characteristics vary widely according to type of laser used. 2. Only suitable for sub-video bandwidth in gas laser. 3. Not suitable for solid state lasers.
Polarization	1. Same considerations as Pockels effect intensity modulation. 2. Twice as much power is transmitted compared to Pockels effect intensity modulation. 3. Improvement in S/N in the presence of background noise over intensity modulation.	1. Similar consideration as Pockels
Frequency using Pockels Cell	1. No transmission loss in system if applied frequency modulation voltage is equal to half wave retardation voltage. 2. Improved S/N over intensity modulation. 3. Large bandwidth capability (Gc).	1. Difficult to detect. 2. Bandwidth limited by longitudinal laser modes. Limits modulation index 3. Would require long lengths of crystal for low modulation power. 4. Requires circularly polarized frequency modulated field. 5. Requires stable laser frequency.
Phase using Pockels Cell	1. No transmission loss in 2. Improved S/N over intensity modulation. 3. Large bandwidth capability (Gc). 4. Requires linearly polarized amplitude mod. field.	1. Difficult to detect. 2. Bandwidth limited by longitudinal laser modes. Modulation index can be larger than frequency modulation. 3. Requires long lengths of crystal for low modulation power.

2.0 MODULATOR COMPONENT EVALUATION

2.1 POCKELS CELLS

As described in Section 1, the electro-optical properties of certain materials can be applied to obtain intensity, polarization, frequency and phase modulation of light. Crystals devoid of a center of symmetry can show a linear electro-optic effect. This is obtained from crystals of two classes: the class 43m of the cubic system and the class 42m of the tetragonal system. Representatives of the former class are cubic zinc sulfide, copper chloride and copper bromide. In the latter class are the dihydrogen phosphates. Table 2-I shows the electro-optic properties of the most important crystals of both classes. The voltage required for half-wave retardation has been calculated for both intensity and phase modulation as described in Sections 1.1.1 and 1.3.1. It is noted that KD_2PO_4 (deuterated potassium dihydrogen phosphate) offers the lowest voltage for both applications. Sphalerite (ZnS) also has good possibilities as a modulator. Although its electro-optical coefficient is low as compared to the dihydrogen phosphate, its index of refraction is much greater, a fact that becomes more significant when raised to the third and fourth powers as is the case in the applications studied. Furthermore, as pointed out earlier cubic crystals like sphalerite can be used with much larger apertures. Unfortunately, to date Z_nS crystals are difficult to obtain. Thus for laser beams in the 0.5 - 1.2 micron spectrum the best choice of Pockels effect modulator seems to be KD_2PO_4. It is, however, by no means certain that there are not other materials, not yet developed, which might offer better properties for light modulation.

2.1.1 Pockels Effect Modulation Up To Video Frequency

At low frequencies most Pockels effect crystals display relatively strong piezoelectric effects. The piezoelectric strains cause additional phase retardation. Above the piezoelectric resonant frequency this effect is lost. The resonant frequency is given approximately by the expression

$$f = \frac{48}{d} \quad kc \qquad (1)$$

where d is the crystal's square edge dimensions in inches. The change in optical retardation before and after resonance is most pronounced in ADP. In KDP having a lower piezoelectric coupling coefficient the effect is much smaller. A one inch square KDP crystal light modulator, for instance, will be resonant at 48 kc showing approximately a 10 percent change in response before and after resonance.

Since relatively large voltages are required to induce a significant optical retardation the upper frequency limit in the sub-microwave region is usually set by heating of the Pockels cell electrodes due to high r. f. currents. To make an approximate estimate of the currents involved, the Pockels cell may be considered a pure capacitance (the loss tangent is <0. 05 percent). If V is the maximum applied voltage, ω_m the modulation frequency and C the capacitance, the RF current, I, that the electrodes are required to withstand is

$$I = \omega_m CV \tag{2}$$

The types of electrodes used may be either simple copper plates with holes to allow passage of light to gold grids or rings designed to minimize fringing of the electric field. The maximum current through the electrodes should be kept smaller than 0. 5 amps. If C is the capacitance in pfds. , V the maximum electrode voltage in kilowatts and f the frequency in megacycles (2) becomes

$$V = \frac{80}{Cf} \tag{3}$$

For a 4 cm^2 $KD_2\,PO_4$ crystal d cm thick the capacitance is

$$C = \frac{15.9}{d}\ pf$$

substituting in (3)

$$V = \frac{5d}{f} \tag{4}$$

Equation (4) is plotted in Figure 2-1. Video frequencies with 100 percent modulation are then possible since in the case of $KD_2\,PO_4$ half wave retardation is obtained with 3. 4 kv. The Pockels effect may be enhanced by stacking a series of Pockels crystals mounted electrically in parallel and oriented in such a manner that their x - y crystal axes at 90^o with respect to one another. When two such crystals are mounted as shown in Figure 2-2 only half of the voltage is necessary to produce the same effect in one crystal. The modulation voltage is reduced by a factor equal to the number of stacked crystals used. The overall capacitance of the system is of course increased by this factor. However, the required input current remains constant.

The fringing of the electric field between the electrodes provides a field in the transparent region. This introduces non-uniform phase retardations

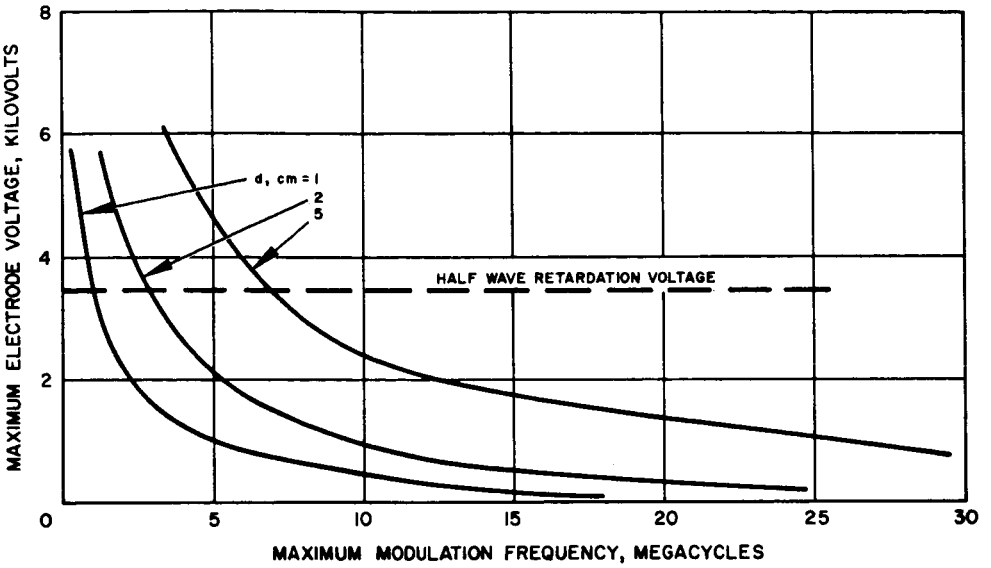

Figure 2-1. Maximum Electrode Voltage at Different Frequencies of A and KD_2PO_4 Pockels Effect Crystal of $4cm^2$ Crossectional Area and d-cms Thick When the Electrode Current is Limited to 0.5 amps.

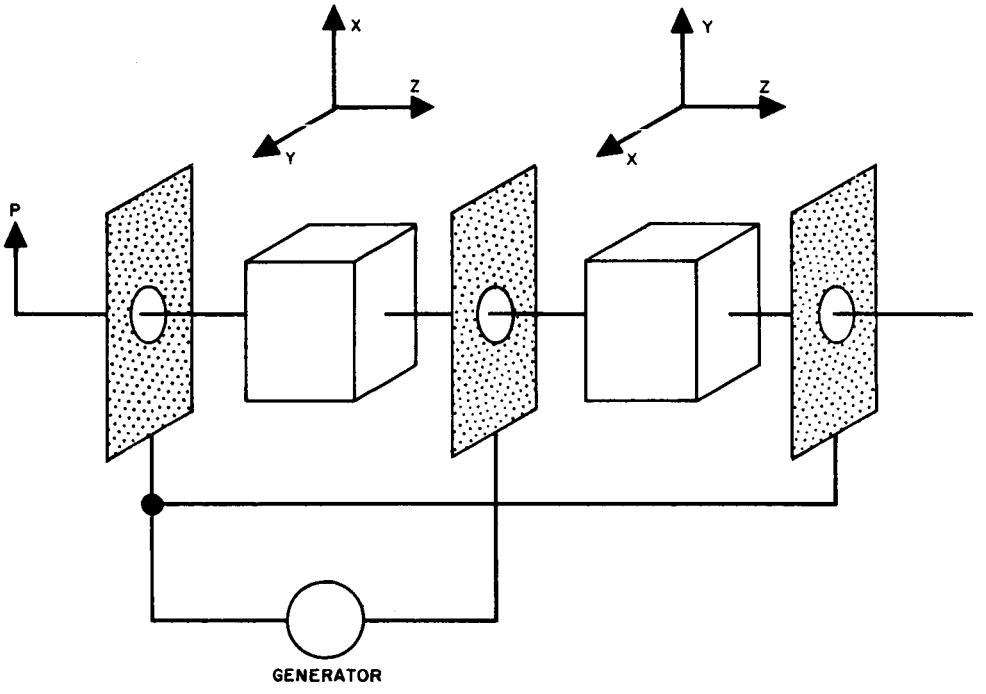

Figure 2-2. Two-Stage Pockels Effect Light Modulator Maximum Electrode V. at Different Frequencies of αKH_2PO_4 Crystal of a $4cm^2$ when the Power Dissipation is Limited to 0.5 Watts

over the aperture and results in a higher voltage than that theoretically calculated to obtain a given response. These effects will not be intollerable if the modulator contrast ratio is not especially sensitive to the specific retardation.

The power dissipated in the crystal itself due to the loss tangent, or dissipation factor, can be theoretically calculated. It is

$$P = \omega_m c V^2 = \omega_m \epsilon'' \epsilon_o \frac{A}{d} V^2 \qquad (5)$$

where the parameters are the same as defined above and ϵ'' is the reactive complex relative permittivity. For a $KH_2 PO_4$ crystal if $A = 4cm^2$, $d = 1cm$, $V = 1$ kv, $\omega_m = 2\pi \times 10^7$ cps, $\epsilon'' = 0.01$ and $\epsilon_o = 8.85 \times 10^{-14}$ fd/cm

$$P = 2\pi \times 10^7 \times 10^{-2} \times 8.85 \times 10^{-14} \times 4 \times 10^6 = 0.222 \text{ watts}$$

a power dissipation which probably can be handled by the crystal. However, power being proportional to the square of voltage an increase in voltage may rapidly surpass the power limit of the modulator. Equation (5) is plotted in Figure 2-3 for several crystal thickness. It is assumed that .5 watts is the maximum power that a crystal of the sizes being considered can dissipate. It is shown that at high frequencies 100 percent modulation will not be possible.

Perhaps the main limitation toward practical use of a Pockels effect modulator at video frequencies is the power supply. It should have the following minimum requirements:

Voltage: 1000 volts
Current: 0.1 amps
Frequency Response: Flat to 2 mc.

Such video power amplifiers have been developed for TV transmitters. The amplifier size is approximately 10" x 20" x 6" and would weigh 2 to 3 lbs. It requires a conventional 1000 and 300 volt DC power supply.

One possible Pockels cell intensity modulator suitable for video modulation could take the following form. To obtain 100 percent intensity modulation with less than 1 kv an arrangement similar to Figure 2-2 would be used. If three such $KD_2 PO_4$ crystals were mounted using an optically biasing quarter wave plate a theoretical peak voltage of only 0.54 kv would be required to obtain 100 percent modulation. Using 2 cm cube crystals at a maximum modulation frequency of 5 mc the electrode current and power dissipation would be kept well below the limit as shown in Figures 2-1 and 2-3. The unit could be mounted within a bath of mineral oil in a quartz container which would facilitate cooling and reduce reflection losses. The light modulator size would be about 2.5 x 2.5 x 10 cm and would weigh about 2/5 pounds.

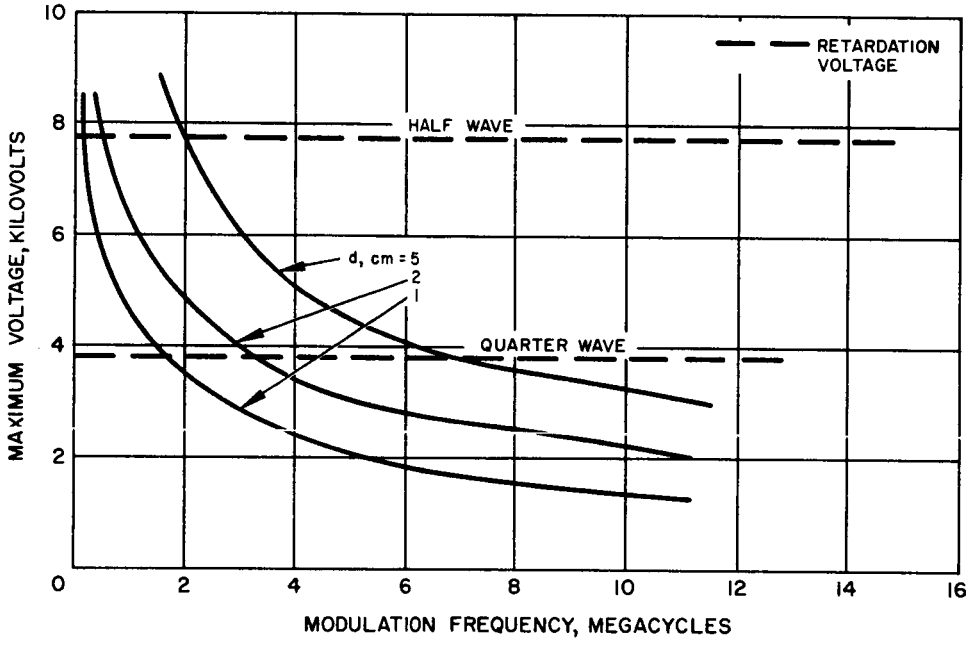

Figure 2-3. Maximum Electrode V at Different Frequencies of a KH_2PO_4 Crystal of a 4 cm^2 When the Power Dissipation is Limited to 0.5 Watts

2. 2 MICROWAVE LIGHT MODULATION

Light can be modulated at microwave frequencies by using the electro-optical effect. Two general approaches may be taken to do this. One consists in building a high-Q structure which would apply large field strength across the crystal with relatively low power inputs. A high Q re-entrant type capacitively loaded coaxial-line resonator has been used to build such a modulator at gc frequencies and few watts input. The high Q requirements of such a modulator make the device inherently narrow band. A second approach consists of making a traveling electric field interact with light propagating through the electro-optical medium. If the velocity of the microwave traveling electric field could be synchronized with that of the light beam the Pockels effect would be enhanced resulting in the possibility of wide-band light modulation and economical use of microwave power. Both approaches will be studied below.

2. 2. 1 Resonant Cavity Modulator

A cross section of a cavity resonator is shown in Figure 2. 4. Its resonant frequency can be determined with reasonable accuracy by setting the susceptance of the capacitance formed by the crystal equal to the inductive reactance of the coaxial transmission line formed by the outer and center conductors. Due to the necessity of boring a hole through the center of the cavity to pass a collimated light beam through, design equations available for the conventional re-entrant cavity can only approximate the actual characteristic of the particular physical configuration employed. The cavity's resonant wavelength, shunt impedance and unloaded Q have been derived (*). The RF losses in the crystal must be taken into account by paralleling the shunt resistance with the dielectric resistance of the crystal R_c where

$$R_c = \frac{1}{\omega_m C \tan \delta}$$

tan δ being the loss tangent. The power dissipated in the crystal is

$$p = \omega_m C V^2 \tan \delta = \omega_m \epsilon'' \epsilon_o \frac{A}{T} V^2$$

A cavity resonant at 1150 mc having a Q = 1300 and tunable between $\pm$ 15mc (using a tuning screw) has been designed. Dimensions are shown in Figure 2-4. The electro-optical crystal used is KDP. Figure 2-5 shows a plot of power input to the cavity and dissipated by the crystal as a function of percent intensity modulation. The high Q of the cavity does not allow broadband modulation. For the particular cavity designed above the bandwidth is

(*) "Microwave Transmission Design Data", T. Moreno, Dover Publications Inc., New York, N. Y., 1948, p 227-240.

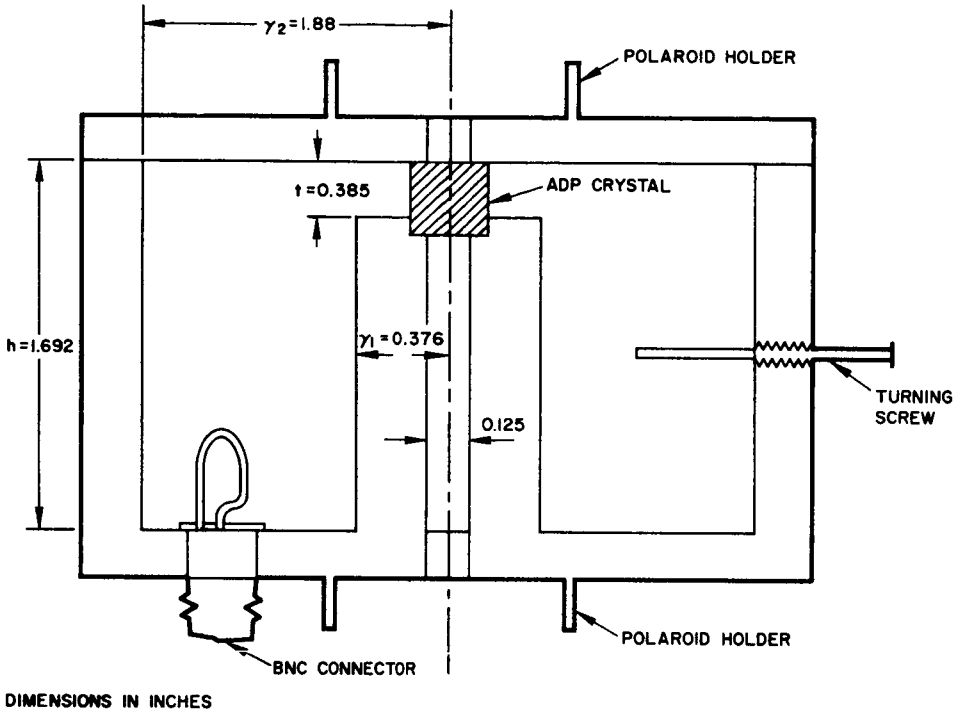

Figure 2-4. Crossection of Reentrant-Type Cavity Used for Microwave Light Modulator

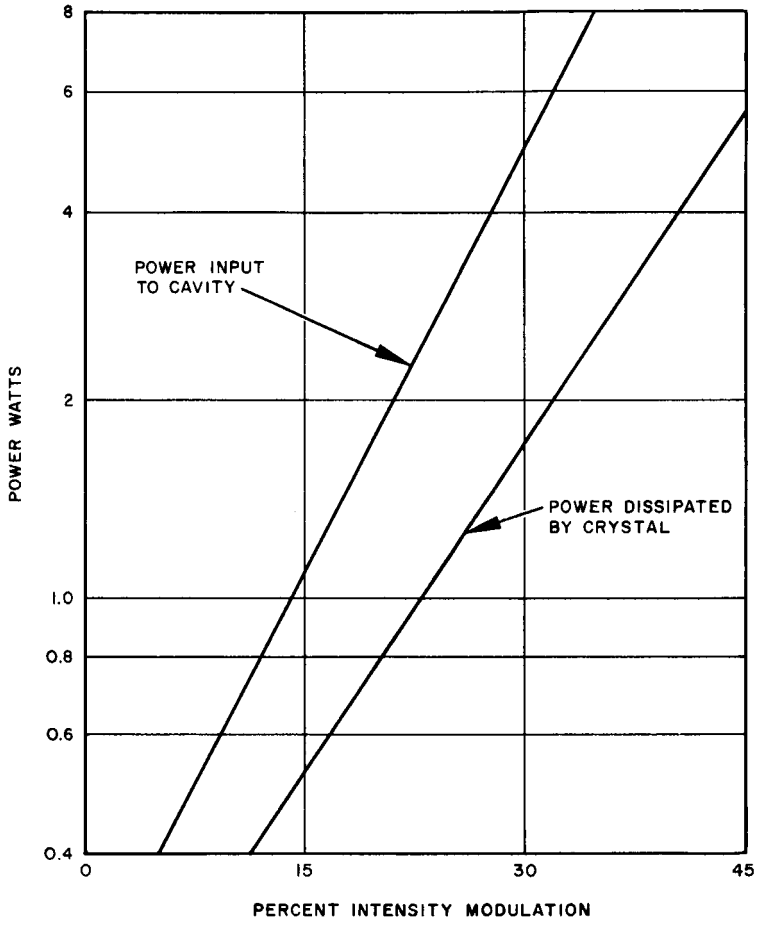

Figure 2-5. Percent Intensity Modulation Versus Power Input and Power Dissipated by Crystal for a Resonant Cavity Light in Intensity Modulation

$$B = \frac{f_m}{Q} = \frac{1150 \times 10^6}{1300} = 0.9 \text{ mc}$$

An increase in bandwidth can be obtained by lowering the Q of the cavity. By so doing, however, a larger power input is required.

The size of the cavity is approximately 30 in.^{3} with a weight of less than 5 lbs. Equipment to produce 10 watts at 1000 mc including power supply and temperature control devices would weigh about 100 lbs.

2.2.2 Traveling Wave Modulator

Wideband light modulation can be afforded by traveling wave interaction in the electro-optic crystal. The Pockels effect is enhanced by making the polarized light beam travel in synchronization with the forward and backward components of the microwave modulative field, in successive passes through the crystal. In a KDP crystal the speed of light is normally three times that of microwaves. To synchronize the speed, a lightly loaded cylindrical cavity containing a Pockels crystal rod of small diameter and excited in the TM_{013} mode may be used. However, the fact that most of the microwave energy is dissipated in the cavity and therefore most of the stored field is located outside the rod make this structure quite inefficient. Figure 2-6 shows a traveling wave modulator of this type. The length of the rod is an integral number of half wavelengths. The longer the length the stronger the electro-optical interaction. When the velocities are matched, the forward wave provides a constant electric field at a given point along the light wave train, while in the backward wave the phase retardation cancels out.

Using a KDP crystal rod 0.4 cm^2 by 3.4 cms long a bandwidth of 5 mc centered at 2700 mc was obtained.* Figure 2-7 shows the microwave power needed as a function of percent intensity modulation. With 10 watts input power over 40 percent intensity modulation can be achieved. The weight of the modulator would not exceed 5 lbs. However, estimated weight of associated equipment including power supplies to produce 10 watts at 2700 mc would run above 200 lbs.

The efficiency of the traveling wave modulator described can be improved if the TEM wave would have most of the electric field in the electro-optical crystal. In such a case the microwave structure could be made essentially dispersionless, a necessary requisite for a wideband light modulator.

It can be shown ** that the conditions for synchronization can be satisfied when the modulating field is propagated as a TEM wave along a plane-parallel guide. The optical z-axis of the crystal is oriented perpendicular to the parallel planes. A traveling wave phase modulator can be built using

(*) "Wide Band Microwave Light Modulator", W.W. Rigrod and I.P. Kaminow, Proc. I.E.E.E., January 1963, p 137.

(**) A. G. Siegman, Applied Physics Letters, November 1962.

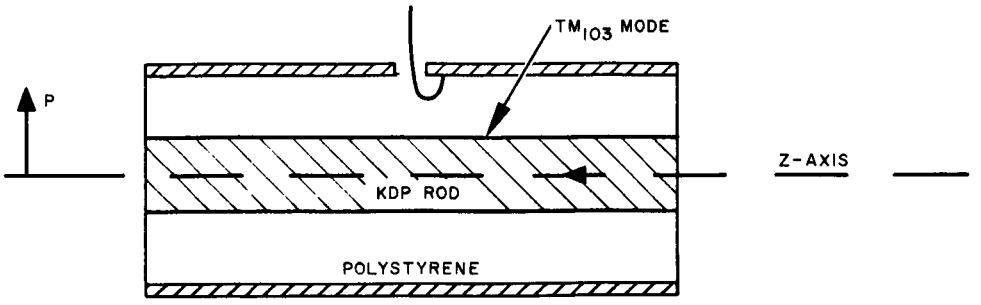

Figure 2-6. Traveling-Wave Modulator

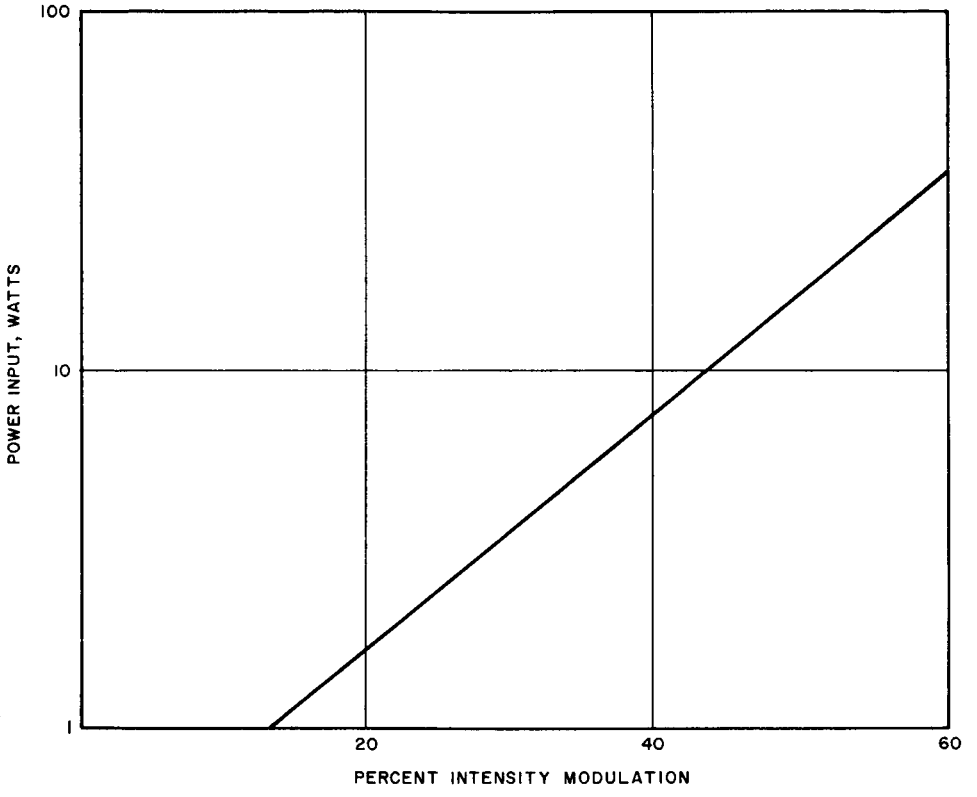

Figure 2-7. Power Input versus Percent Intensity Modulation for a Traveling-Wave Light Modulator

0.5 megacycle bandwidth

the principles studied in Section 1.3.1. The cross sectional view of a phase modulator is shown in Figure 2.8. The parameters are calculated from the following relationships (*)

$$a = \frac{\omega(\epsilon_1 - 1)}{n^2 - 1} \tag{1}$$

and

$$b = Z_o \frac{\left[a(a - \omega + \epsilon_1 \omega)\right]^{1/2}}{120\pi} \tag{2}$$

where

ϵ_1 = the relative dielectric constant of the crystal

n = the index of refraction

Z_o = the characteristic impedance of the transmission line

The length of the crystal, ℓ, is directly proportional to the phase retardation, δ. As was noted in Section 1.3.1 the relationshis is as follows

$$\ell = \frac{\delta\lambda}{\pi n^4 r E_Z} \tag{3}$$

Also, it may be observed that the longer the crystal the smaller the field strength required for a given phase retardation.

Figure 2-9 shows a plot of length of crystals required as a function of electric field across the transmission line planes for various modulation indices. Assuming the transmission line to be terminated by a matched load the power required to support these fields is approximately equal to

$$P_{in} = \frac{E_z^2 b^2}{Z_o} \tag{4}$$

where the attenuation introduced by the electro-optical crystal is being neglected.

(*) "Traveling Wave Phase Modulator", C. J. Peters, I. E. E. E., January 1963, p. 47.

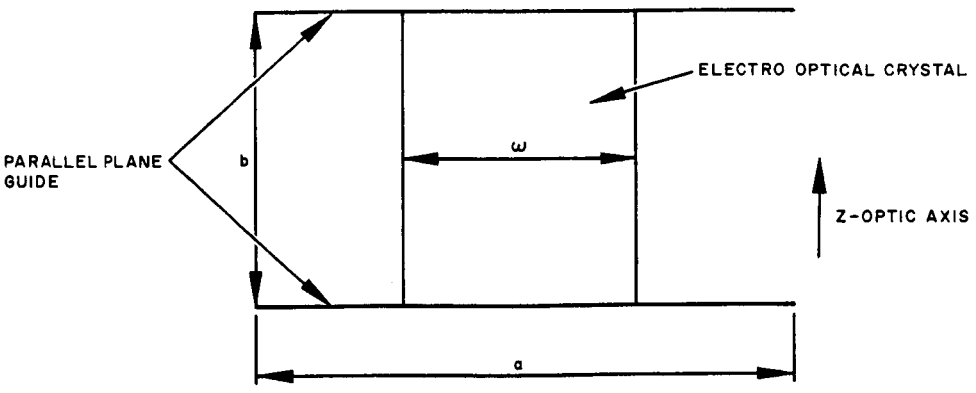

Figure 2-8. Crossection of Traveling-Wave Phase Modulator

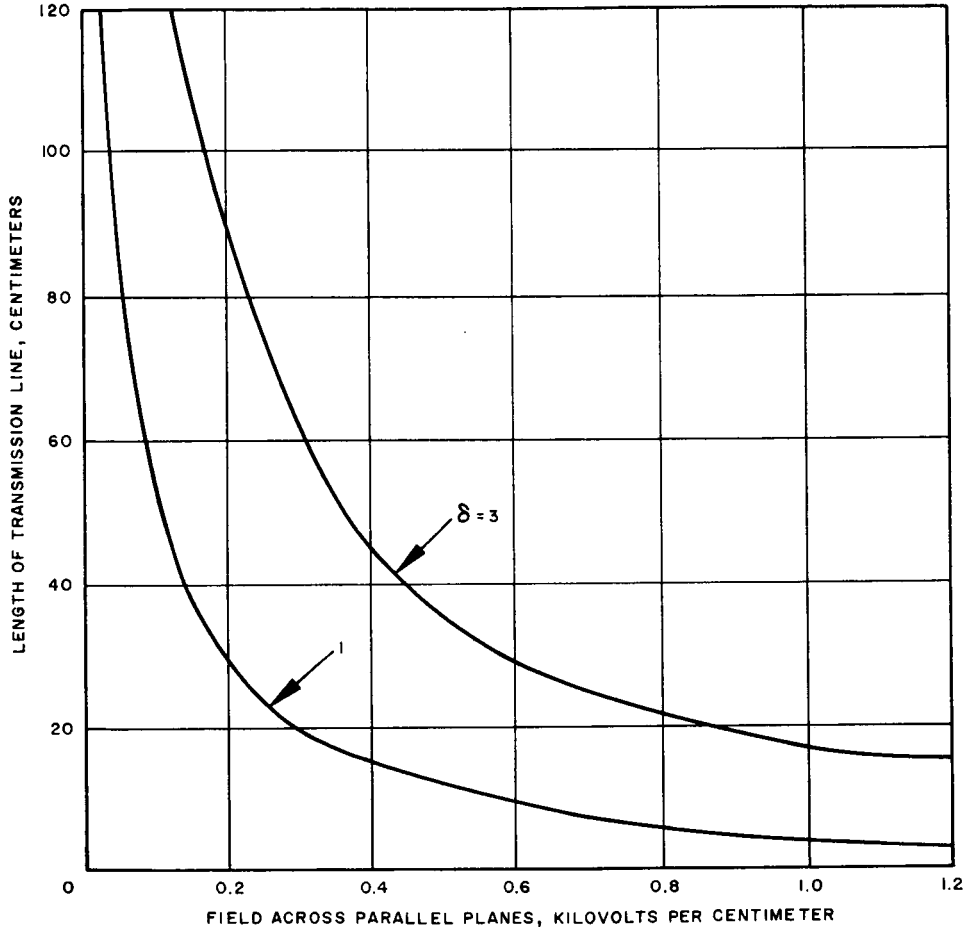

Figure 2-9. Transmission Line Length as a Function of Field Across the Optic Axis for a Modulation Index $\delta = 1$ and $\delta = 3$

This is justifiable if

$$Z_o \omega \epsilon'' \epsilon_o \frac{\omega \ell}{b} << 1. \qquad (5)$$

For $\ell = 100$ cms, $Z_o = 50$, $\lambda = 0.5\ \mu$, and using a KDP crystal as a phase modulator the power requirements as a function of modulation index are shown in Figure 2-10. The power input is relatively low if the crystal aperture is small. Over the 100 cm length the transmission loss of light through KDP should be less than 6 db. Since the aperture of the modulation is probably smaller than the laser beam width a recollimation of the light would be necessary.

To obtain phase modulation over wide bandwidths with relatively low powers a long modulator is necessary. The size of the modulator described would be approximately 2 x 2 x 100 cms. and would weigh about 10 lbs. The same modulator can be used at low or high frequencies. At 100 mc about 30 lbs. of equipment are necessary to produce an output of 10 watts.

The length of the crystal may be reduced by having the laser beam zig-zag across the microwave beam. (Figure 2-11) The angle of incidence is critical in establishing syncronization between the microwave and light beam velocities. To obtain intensity modulation the incident light would have components of polarization in both principal planes of the crystal. In this case the neutral bi-refringence of crystals, like the hydrogen phosphates, limits the optical pathlength severely since the beam is not along the optic axis. In cubic crystals such as ZnS this limitation does not exist.

2.3 SUMMARY

Table 2-II summarizes the properties of Pockels effect modulators. For narrow bandwidths, modulation at microwave frequencies (subcarrier) offers the advantage of low power requirements as compared with video frequency modulation and avoids the possibility of low frequency distortion. However, these advantages should be compared with the increased complexity of using microwave systems both to modulate the Pockels crystal at the transmitter and to detect it at the receiver. Sections 4 and 5 discuss tradeoffs between the different types of detection and demodulation systems.

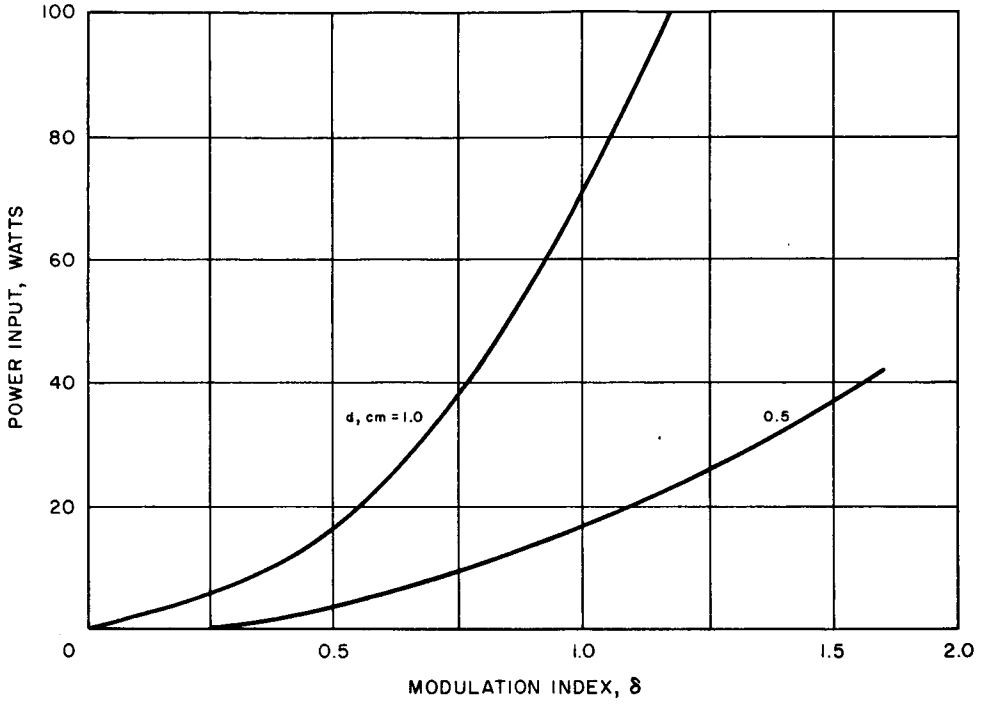

Figure 2-10. Power Input to Phase Modulator as a Function of Modulation Index (l = 100 cms) for b = 1 and b = 1/2 cms

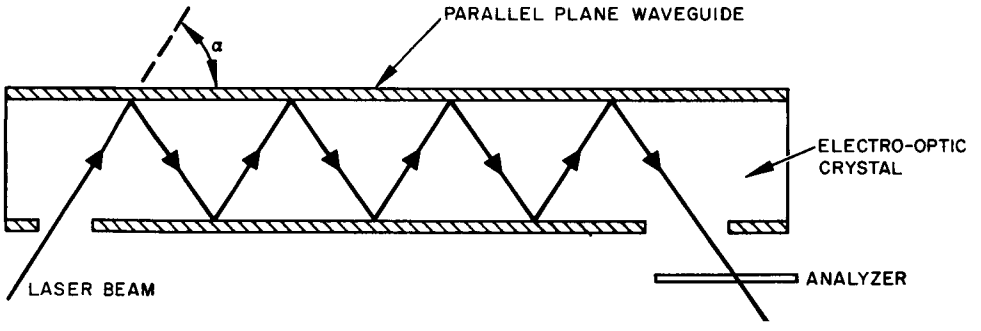

Figure 2-11. Traveling-Wave Light Modulator

TABLE 2-II

CHARACTERISTICS OF POCKELS EFFECT MODULATORS

MODULATION FREQUENCY	PROPERTIES
Video	1. $KD_2 PO_4$ offers the best electro-optical characteristics. 2. Low frequency response is affected by piezo-electric effects. 3. Capacitance of modulator induces high reactive currents at upper frequencies. 4. Electrode current capacity and crystal power dissipation limit upper frequency response. 5. A 2 cm^2 aperture modulator is designed for 0-5 mc response requiring 0.54 kv for 100 percent modulation, and 100 watt power supply.
Microwave	1. Requires lower power than video modulation. 2. Wide bandwidths require very long devices for low input modulation power. 3. A cavity type modulator requires 5 watts for 30 percent modulation at 1 gc and 1 mc bandwidth. Length of device 6 cms. Area 0.4 cm^2. 4. A traveling wave modulator (narrowband) requires 10 watts for 40 percent modulation at 3 gc and 5 mc bandwidth. Length 5 cms area 0.4 cm^2. 5. A traveling wave modulator (wideband) requires 10 watts for 100 percent modulation and 3 gc bandwidth. Length 100 cms area 0.25 cm^2.

TABLE 2.1 PROPERTIES OF ELECTRO-OPTIC CRYSTALS

	KD_2PO_4	KH_2PO_4	$NH_4H_2PO_4$	RbH_2PO_4	KH_2AsO_4	ZnS	CuCl
Electro-optical constants (10^{-7} cm/kvolt)	r_{63}: 23.6 r_{41}: -	10 8.6	8.5 28	11 -	11.6 -	- 2	- 1.8
Half-wave* voltage (kv)	3.4	7.8	9.6	7.3	6.2	8	18
Index of refraction		1.509	1.524	-	1.567	2.368	1.973
Dielectric constant	45	21	15.4	15	21	8.4	
Transmission range (micron)	0.4-1.2	0.4-1.2	0.4-12	-	-	-	0.4-15
Half wave (**) voltage (kv)	2.3	5.2	6.3	4.9	4.0	3.4	9.1

* Voltage necessary to produce half-wave phase retardation when light and electric field are parallel to the optic axis ($\lambda = 0.5$ microns)

** Voltage necessary to produce half-wave phase retardation when light travels parallel to the y axis, is polarized along x axis and the E field is parallel to the z axis.

3.0 MODULATION SYSTEMS

The initial inputs from the television camera, the voice channel, and the ten telemetry channels must be processed in some manner before it can be applied to the modulator. For the purposes of this section it is assumed that system considerations dictate a relatively narrow information bandwidth, so that particular mention is made of bandwidth reduction schemes. The various inputs are combined by either frequency or time multiplexing and are then amplified to a level sufficient to drive the modulator.

Some modulating devices which are practical for light beam modulation were discussed in the previous section. The question of the form of the modulation signal will be considered on the following paragraphs; specifically, the relative advantages and disadvantages of analog and digital modulation systems will be discussed.

3.1 ANALOG SYSTEMS

The most direct and simplest method of information transmission is by analog modulation. The modulating signal is used in its original form to amplitude or angle modulate the CW laser or to pulse position or pulse amplitude modulate the pulsed laser. The modulating signal requires no processing or alteration and the bandwidth required is the same as that of the signal. The characteristics of various kinds of analog modulation have been thoroughly defined and explored. In general, the characteristics include narrow unprocessed bandwidth, a relatively high susceptibility to noise interference, and a high rate of redundancy.

Multiplexing of analog information signals usually requires a frequency division process whereby each signal is on a different subcarrier frequency all of which are combined to form the modulating signal. If memory storage is available along with capabilities of differential read in and read out times, time sharing multiplex, where each information signal is allotted a period of time to modulate the common carrier may be used. For comparison sake, consider the information channels and converter diagramed in Fig. 3.1. The information channels consist of 10 telemetry channels, each 1000 cps wide, one voice channel, 1500 cps wide and one television channel, which we will assume is 6 kc wide. All of these signals must be combined into one signal which is used to drive the laser modulator.

In the analog system the converter would be a series of subcarrier oscillators and/or mixers if the modulation signal is to be frequency

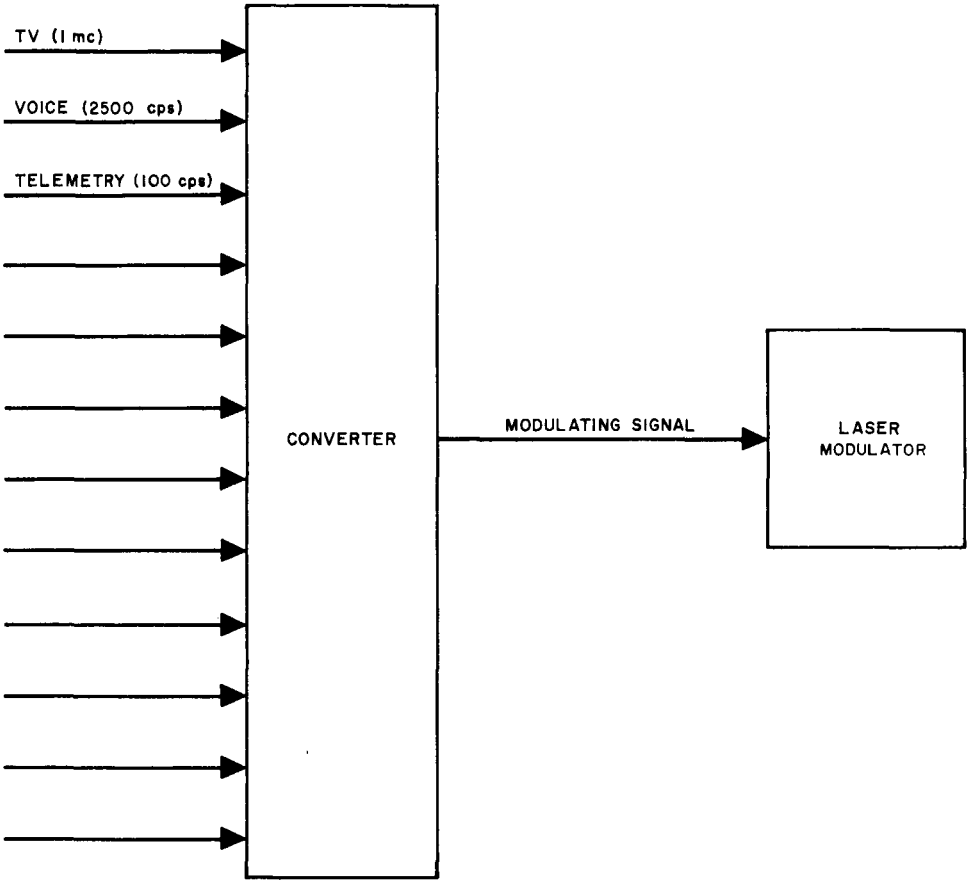

Figure 3-1. Basic Transmission Data Processing System

multiplexed or a timing clock and selector switch, along with memory devices, if time sharing is to be used. With analog information, frequency division requires the least equipment and is most often used. The total bandwidth required, if it is assumed that each signal contains sufficient guard band is only slightly over 17 kilocycles. Analog systems to not particularly lend themselves to a redundancy reduction so the necessary modulator bandwidth would be about 17 kc.

A typical converter of this type might occupy 0.7 cu. ft., weigh 4 lbs., and require 5 watts of power to operate.

3.2 DIGITAL SYSTEMS

Digital modulation requires that the incoming information signals be quantized into predetermined energy levels. Some means is then selected to identify the particular level. The received signal is then reconstructed by using this information and filtered to restore the original signal. The difference between the analog signal and the quantized signal is known as quantization noise and is inherent in the system. However, this noise may be reduced to tolerable level, by judicious selection of the quantization and encoding process.

Two digital systems are in general use: pulse code modulation (PCM) and delta modulation (DM), or a combination of the two (DM-PCM). PCM involves quantizing the signal with reference to a base line, usually the lowest value of the signal, within some sampling period, usually selected to be about equal to the fastest rise time expected. If the number of quantization levels is chosen as eight, and a binary code is to be used, then 3 bits are required to determine the level in each sampling period. If the highest information frequency is 6 kc and the Nyquist criterion is used, a minimum 12 thousand sampling periods per second are required. With 3 bits per period 36 kilobits/second are required. If the information bandwidth is twice the RF bandwidth, then an 18 kc bandwidth is required. A faster sampling rate and closer quantization levels contribute to higher fidelity and greater bandwidth requirements.

DM, instead of using a base reference as starting point for each level, uses the previously transmitted level as a starting point and then goes up or down 1 step, depending on the relative analog signal level. A greater number of sampling periods may be used since each pulse transmits information about the signal. However, sample timing becomes rather important, for a slowly varying low amplitude, analog signal will cause to be transmitted a series of positive and negative pulses. On the other hand, if the sampling periods are too long, the system will not be able to keep up with a rapidly changing signal. Good signal fidelity requires high sampling rate in conjunction with small level changes.

A combination of these two systems, DM-PCM, is frequently used where the reference is taken as the previous level but the magnitude of the change

is coded. In this case, one bit must be used to indicate up or down while the remainder indicate magnitude of the change. For a given bandwidth, DM-PCM transmits less information and has greater fidelity than DM alone, and transmits more information with greater fidelity than PCM alone.

Digital systems, in general, operate with a S/N ratio higher than analog systems in a high noise environment. The uncompressed signal of the digital system requires a wider bandwidth than the analog signal. For instance, considering the system of Fig. 3-1 with 4-bit coding, a bandwidth of greater than 240 kc is required as compared to 60 kc for the analog system. The digital system also requires accurate clocks at both the transmitter and receiver to synchronize the sampling periods. The same multiplexing techniques applicable to analog systems are also suitable for digital systems, although digital systems are more suitable for time multiplexing.

Most information signals contain a high percentage of redundant material and digital systems are especially suited for removing this material. However, it is well to remember that the more sophisticated a system becomes in terms of information transfer the more complex it becomes in terms of components. Therefore a practical limit is imposed on the degree of sophistication of any system. Digital transmission of a speech signal requires about 20,000 bits/second while the basic information content of the signal is about 60 bits/second. Bandwidth reduction on normal TV has been achieved (*) and special, slow scan, high resolution video camera signal has been compressed into a 6 kc bandwidth (**). By using compression techniques digital signals can make better use of available bandwidth with attendant increases in S/N ratio than do analog systems.

3.3 TYPICAL SYSTEMS

A block diagram of a typical analog converter using frequency division multiplexing is shown in 3-2, and that of a typical digital converter in Fig. 3-3. The signal inputs in Fig. 3-1 modulate subcarrier oscillators which transfer the signal to the desired portion of the frequency spectrum. The output signals of the modulated subcarriers are then added together and used to amplitude, phase, polarization or frequency modulate the light beam.

Fig. 3-3 is similar but after the modulated subcarrier oscillators come a quantizer and encoder. The outputs of the encoder which are digital signals, are now added together and again used to modulate the light beam. Now the modulation is confined to two discrete levels. Using a 4 level DM-PCM coding, 3 bits per sample are required. In this system,

---

* W. Schreiber, C. F. Knapp and N. D. Kay, "Synthetic Highs-An Experimental TV Bandwidth Reduction System", 84th SMPTE Convention, Detroit, Mich., Oct., 1958.

** John R. Scull, "A System for Lunar Photography and Data Transmission", Jet Propulsion Laboratory, May 28, 1960.

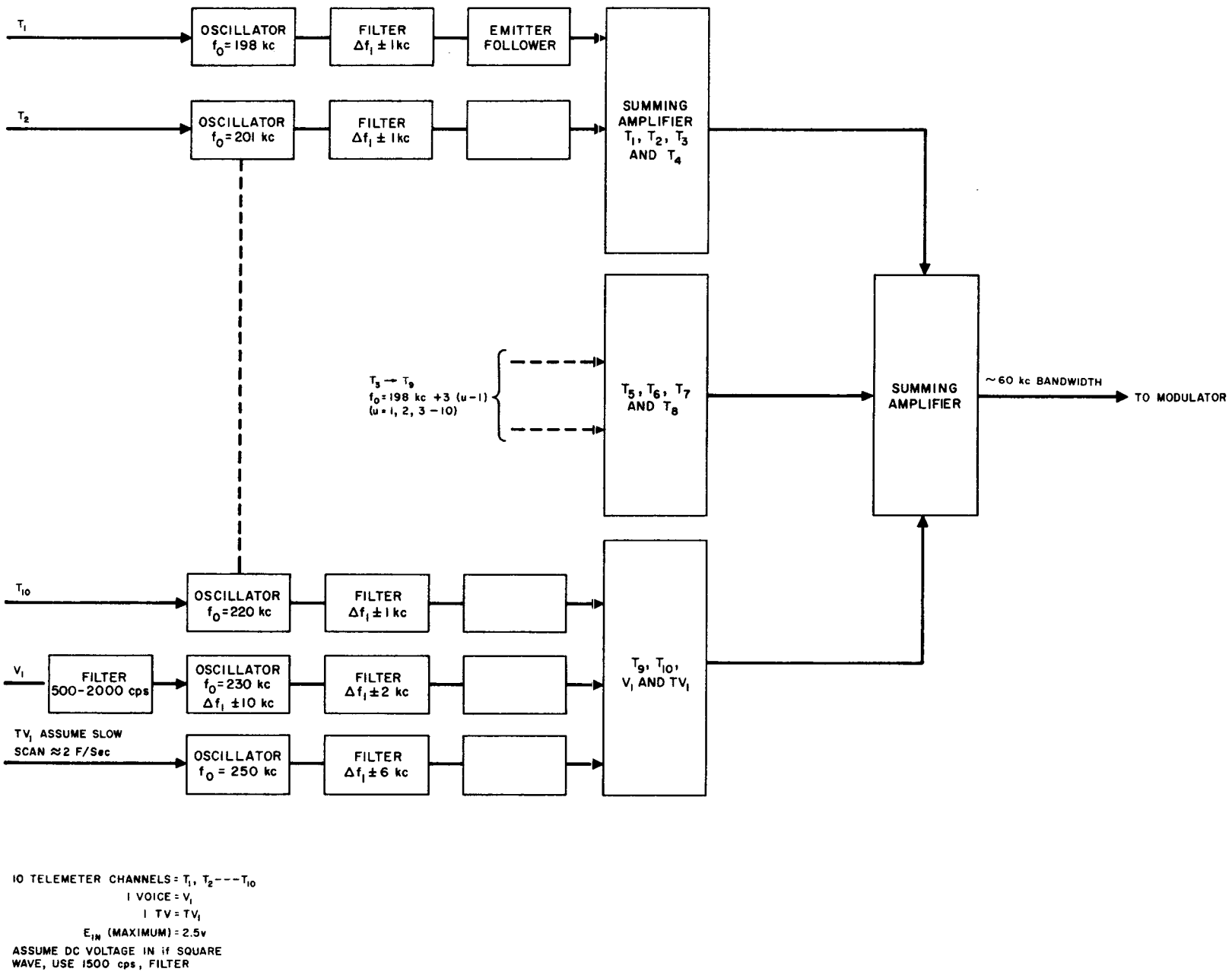

Figure 3-2. Analog Converter

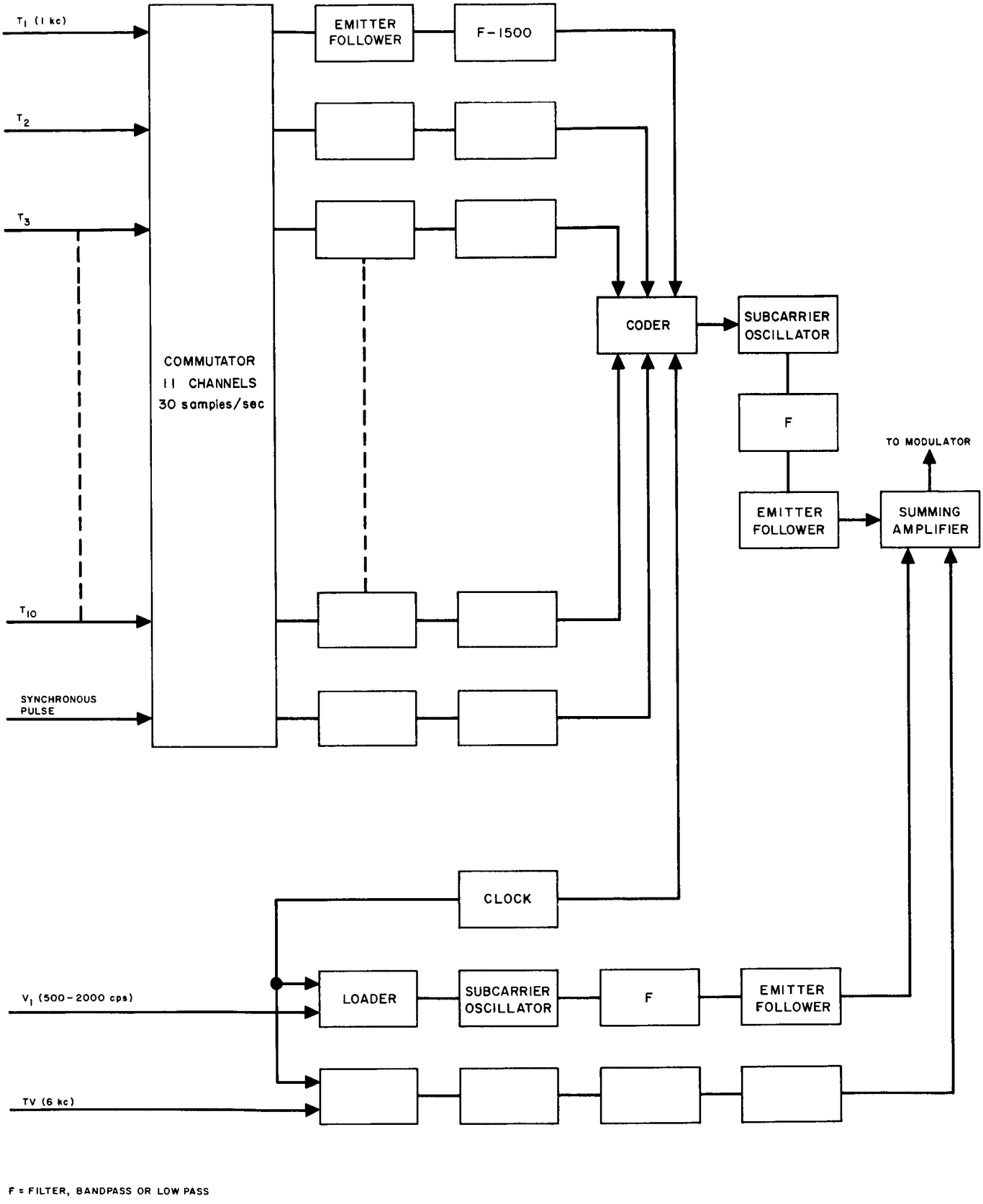

Figure 3-3. Digital Converter

commutated telemetry channels are used, saving nine subcarrier oscillators but at a slight degradation of information capacity. The nature of the information to be transmitted would determine whether or not this would be practical. Although many variations enter into a system such as this, a typical system might weigh 4 lbs., require 11 watts and displace 0.7 cu. ft.

4.0 DEMODULATION COMPONENT EVALUATION

4.1 THE PN JUNCTION DETECTOR

Semiconductors may be used as photodetectors by taking advantage of one of several effects that may occur when light strikes their surface. According to the effect used they are broadly classified as photovoltaic (PV), photoelectromagnetic (PEM), and photoconducting (PC) detectors. In the case of PV detectors absorbed photons produce electron-hole pairs separated by an internal electric field at the semiconductor junction and thus establishing a potential difference. PEM detectors operate on the principle that the electron-hole pairs generated by the photons are separated by an external magnetic field, establishing in this manner a potential difference. In PC detectors the generated free electrons change the resistivity of the material. PN junction photo-diodes fall under this classification and offer the best possibilities for broad-band optical detection. PEM detectors are easier to fabricate than PN junctions, however their power conversion efficiency is low and generally must be cooled to liquid nitrogen temperatures to obtain good sensitivity. At optical wavelengths above 1.6 microns it would be necessary to use a PEM photodetector (such as In Sb), however since the laser frequencies most likely to be incorporated in the optical communication system lie in the visible or near infrared region we shall limit our study to PN junction diodes.

Fig. 4-1 shows the basic configuration of a PN junction detector. When light of a wavelength shorter than the absorption edge strikes the surface of the back biased detector, a hole-electron pair is formed for each absorbed photon. The transit time of the carriers at the junction and the junction capacitance will essentially determine the frequency response limitation of this detector. The junction width is directly proportional to the transit time and inversely proportional to the capacitance.

$$t = \frac{v_o}{\omega_m}$$

where v_o is the saturation drift velocity of the carriers and ω_m is the frequency which the detector must follow. v_o is a function of the absorption constant of the junction, and the diffusion constant for electrons.

The saturation drift velocity of the junction is directly proportional to the area of the junction and establishes its maximum value. A design limit for the junction area A is (*)

---

* H.S. Sommers, "Demodulation of Optical Signals with Semiconductors", Proc. I.E.E.E., January 63, p. 142.

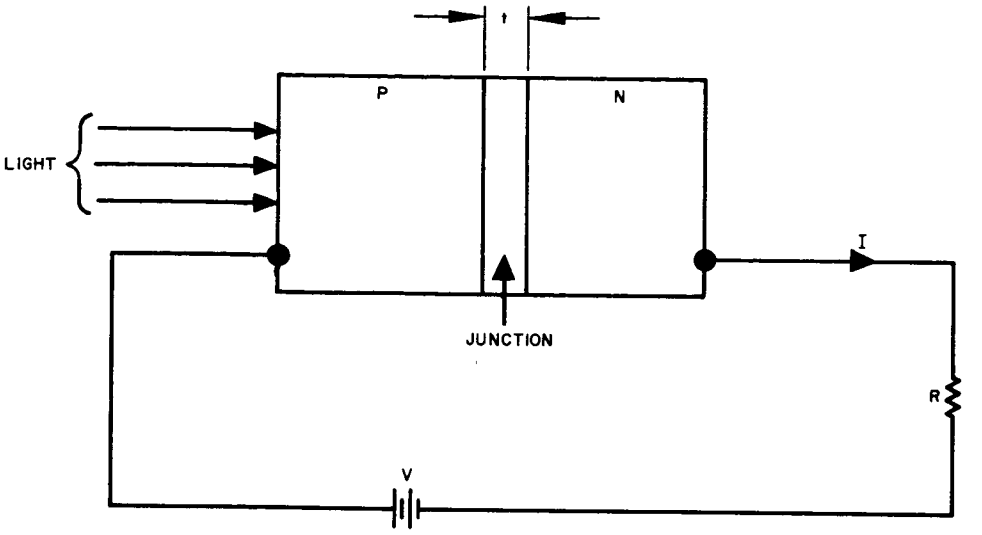

Figure 4-1. PN Junction Detector

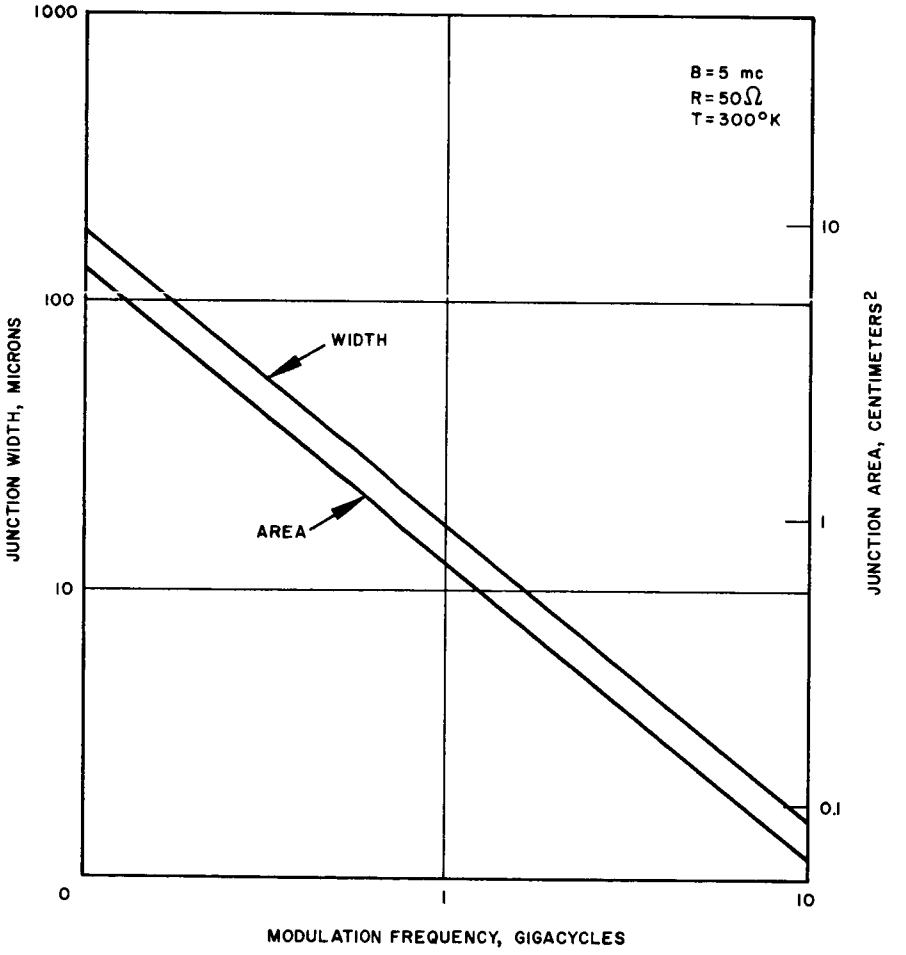

Figure 4-2. Junction Diode Design

$$A \leq \frac{2 \times 10^{12} v_o}{R \omega_m B}$$

where R is the load resistance and B is the modulation bandwidth.

The load resistance also is an important factor in determining the bandwidth limit of the photodiode, since

$$RC \leq \frac{1}{2 \pi B}$$

where C is the junction capacitance. For maximum power output the load resistance must be matched. At microwave frequencies the load resistance is usually the transmission line characteristic impedance (50 ohms). The effect of the internal resistance of the diode is negligible when compared with R.

Fig. 4-2 shows typical design parameters for a germanium or a silicon PN junction photodiode operating at room temperature. The calculations have been made assuming a modulation bandwidth of 5 mc and a load resistance of 50 ohms. The quantum efficiencies of these detectors will be above 0.7 when the radiation wavelength does not exceed 1.5 microns in the case of germanium and 0.85 microns in the case of silicon.

The current available, I, from the photodiode depends on the number of photons/second n striking the diode surface and on the quantum efficiency

$$\eta \ , \ I = \eta n \epsilon .$$

The power P, delivered to the matched load R will be

$$P = (\eta n \epsilon)^2 \frac{R}{4}$$

The minimum detectable signal will depend on the internal and external noise sources of the detector. Assuming background noise to be negligible, there will be three notable noise currents: I_D, the noise from the detector itself mostly in the form of dark current; I_N, the quantum noise of the light beam; and I_R, the thermal noise in the load resistance.

At low frequencies (5 MC) the noise-equivalent power (NEP, the power necessary to produce a signal to noise ratio equal to unity) may be computed from

$$NEP = \frac{(AB)^{1/2}}{D^*} \text{ watts}$$

---

* H.S. Sommers, Ibid.

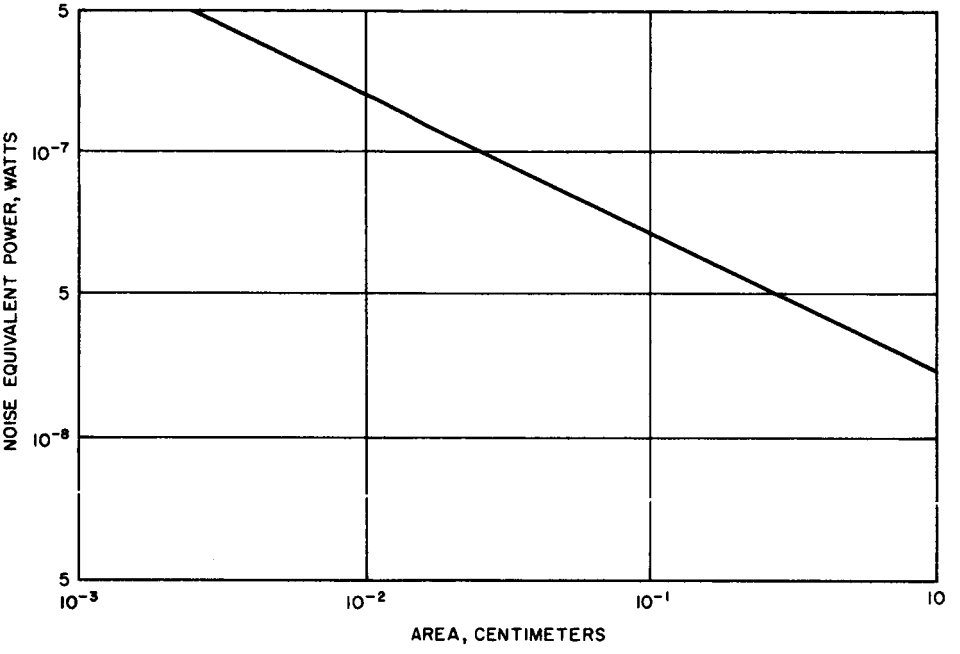

Figure 4-3. Typical Values of NEP as a Function of Detector Area

where D^* is the power detectivity. Taking $D^* = 10$ cm $cps^{1/2}$ ω^{-1} and B = 4 mc, the NEP is plotted as a function of area in Fig. 4-3.

At high frequencies of modulation I_D can be neglected as compared with I_N. In this case the theoretical noise equivalent power can be shown to be (*)

$$NEP \geq \frac{4f}{\eta \lambda} + 10^{-23} \text{ watts}$$

The size and weight of the photodiode itself is negligible. To detect microwave frequency modulated light a high-Q resonant structure must be used to enhance the expected low level signal. The Q must not be high enough, however to degrade the information bandwidth. One such structure is illustrated in Fig. 4-4. The diode is placed in the inner conductor of a coaxial line by using a metalic holder. One end is connected to a microwave receiving system via a movable short tuner. The other end is also connected to an additional movable short. D. C. isolation is provided by a high capacitance break in the metallic holder. A conventional superheterodyne receiver may be used to retrieve the video modulation.

Detection and amplification of the light beam may be obtained by utilizing a parametric amplifier structure. Figure 4-5 shows such a structure. A gain of 15 db has been obtained (*) detecting modulated light at 4 gc when pumping power at 8 gc is supplied. For pump power of about 100 mw a gain of 30 db can be expected with 5 mc modulation bandwidths. Such a parametric amplifier would weigh less than 10 lbs.

4.2 PHOTOEMISSIVE DETECTORS

Phototubes, photomultipliers, and traveling-wave microwave phototubes are of this type. They are the most sensitive and fastest detectors in the ultraviolet and light region of the spectrum.

The operation of a photoemissive detector is based on photoelectric emission, which is also called external photoeffect. This effect consists of two steps: absorption of light by a solid and emission of an electron from the solid. Light may be absorbed without an electron being emitted; this occurs every time an electron-hole pair is created by the absorption of light within a semiconductor.

The basic law of photoelectric effect is that the emission of each electron is caused by the absorption of a single photon. The kinetic energy of the electron leaving the emitting surface is, according to Einstein's formula,

---

* S. Saito, et al., Proc. IRE, November 1962, p. 2369.

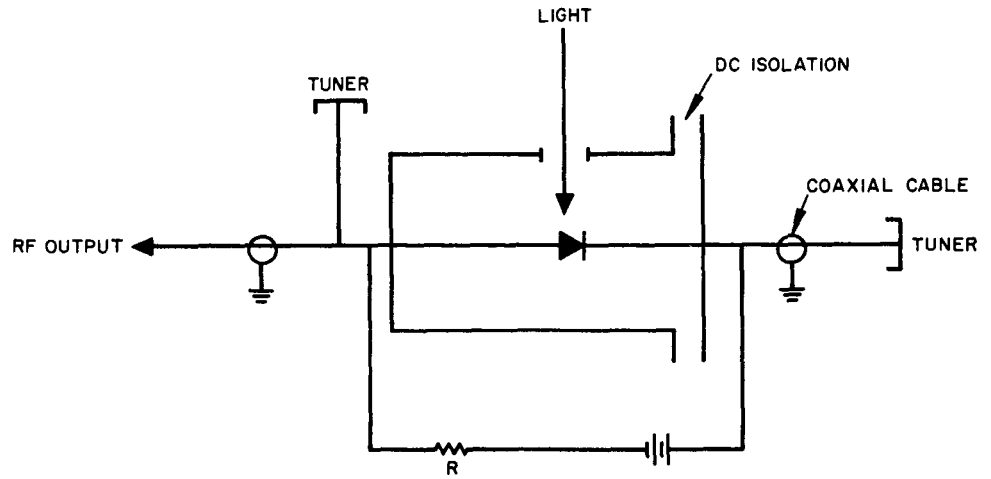

Figure 4-4. Microwave Photo-Diode Circuit

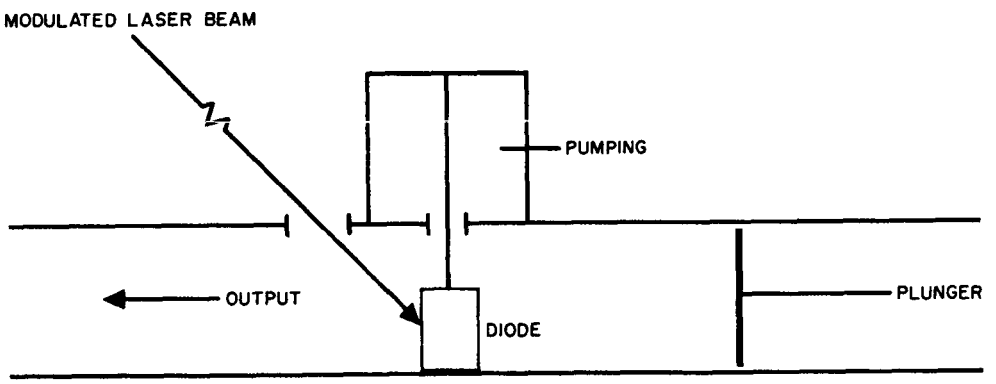

Figure 4-5. Microwave Photodetector Using the Principle of Parametric Pumping

$$E_{kin} = hf - e\Phi$$

where Φ is the work function of the surface. Only light of such frequency can cause photoeffect for which $hf \geq e\Phi$, or

$$\lambda_L = \frac{12,400}{e\Phi}.$$

The element with the lowest work function is cesium. For this element $e\Phi = 1.9$ eV; therefore, the wavelength limit of a cesium photocathode is about 6500 Å. Composite photocathodes consisting of combinations of metals and oxides have lower work functions and are capable of emitting electrons when irradiated with light whose wavelength is longer than 6500 Å. The spectral response of such photocathodes is generally quite complicated. It is characterized by the photocurrent emitted per unit incident radiant power. The spectral response curves are so drawn that their peak is 1 or 100; they represent relative values. The ordinates of these curves represent the electron current emitted by the photosurface upon irradiation with light of wavelength λ in terms of the current that would be produced if the wavelength of the light were changed to that of the most effective radiation for the surface and the intensity of the irradiation remained the same.

The commercially available photodetectors are standardized according to their spectral response. Detectors with S-10 and S-20 responses are most suitable for the detection of radiation in the wavelength region between 6900 and 7000 Å; i.e., for the ruby laser emission at 6943 Å. For the neodymium glass laser (1.06 μ) the S-1 response is best suited. The responses of these devices are shown in Fig. 4-6. It is clear from the figures that the sensitivities of both the S-10 and S-20 surfaces are low at 6943 A compared with their maximum sensitivities which are attained between 4200 and 4600 Å. It should be noted, however, that the quantum efficiency of the S-20 surface at 6943 A (2.5 percent) is roughly 60 times the quantum efficiency of the S-1 surface at 1.06 μ. For this reason, given equal transmitter power and beam collimation, the ruby system is to be preferred over the neodymium-glass system because of the higher available detector sensitivity.

A figure of merit for a photoemissive surface is the quantum efficiency, the number of emitted electrons per incident photon. The quantum efficiency, q, the incident power, W, and the photoelectric electron current are related as follows: The number of incident quanta per second is

$$n = W/hf,$$

the electron current emitted from the photosurface is qne, and, therefore,

$$i = qWe/hf$$

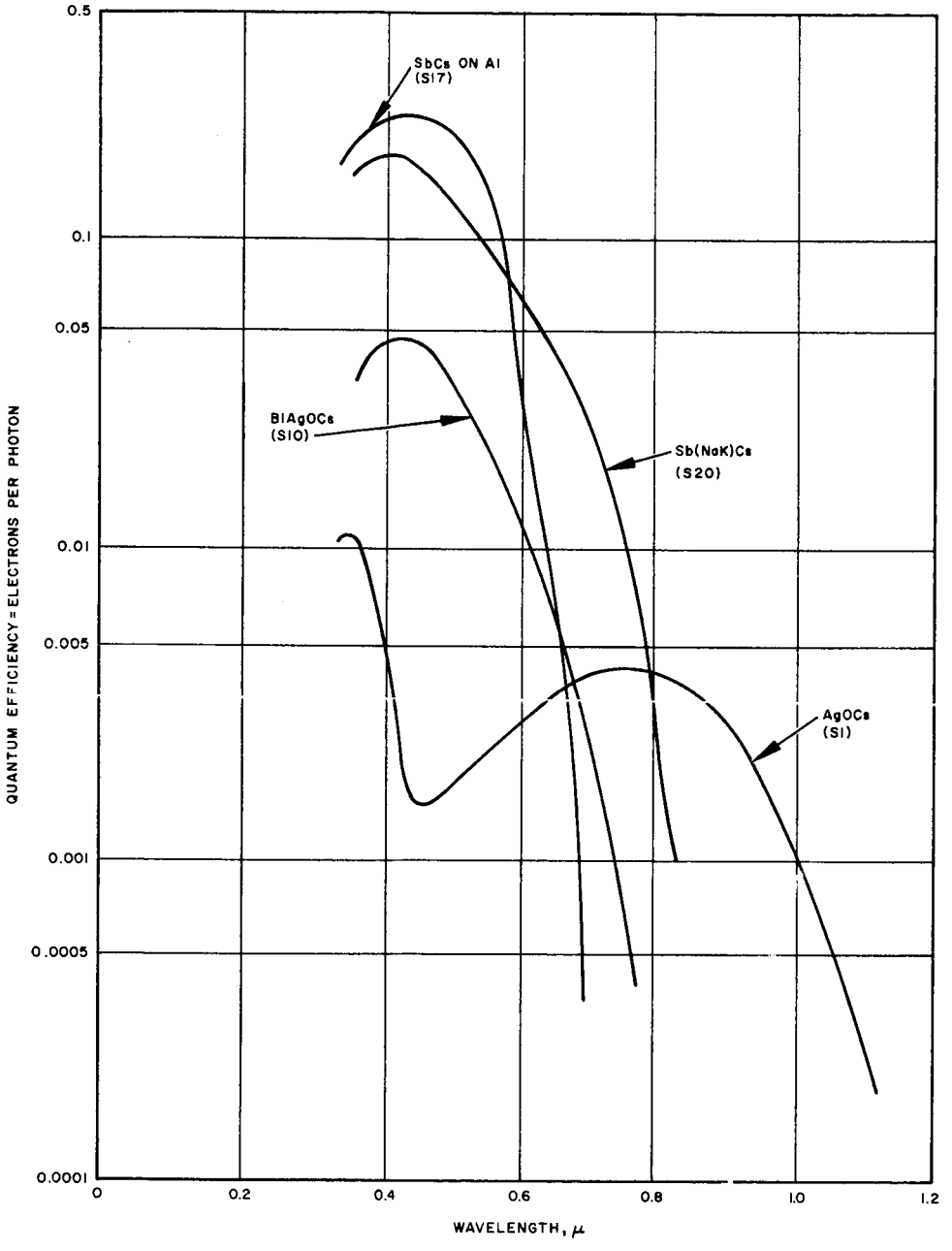

Figure 4-6. Spectral Response of Various Photo Cathodes in Visible and Near-IR Region

The quantum efficiency is not the only figure of merit of a photodetector. Since each detector generates a certain amount of noise the following question is relevant:

What is the minimum intensity of radiant power falling on the detector which gives rise to a signal voltage equal to the noise voltage from the detector?

This question can be answered only if the spectral composition of the incident radiation is specified. Even when this is done the answer is not unique. The measured noise voltage in the output of the detector is proportional to the square root of the electrical bandwidth; therefore, the electrical frequency intervals must be specified. Some ambiguity still remains because the noise from the detector is temperature dependent. In fact, a detector sensitive in the orange-red region is quite noisy when operated at room temperature. For high sensitivity it should be refrigerated.

A figure of merit, the Noise Equivalent Power (NEP, is defined as follows: it is the radiant power falling upon the detector capable of producing an output signal equal to the output due to (internal) noise alone. In other words, NEP is the signal for which the s/n ratio is unity. The quality of incident radiation involved in the definition of NEP is specified either by the temperature of the black-body radiation which serves as a reference source, or in terms of the wavelength if a monochromatic source serves as a reference. The NEP is measured with a chopped incident radiation and the chopping frequency is usually chosen in the low audio frequency range. The noise is always referred to a 1-cps bandwidth of the output.

The reciprocal of the NEP is the detectivity, D. This quantity expresses the s/n ratio obtainable/watt of incident power. A higher value of detectivity indicates a better detector, a detector capable of detecting a signal of a smaller intensity.

The sensitivity of a photodetector is the ratio of output current to incident radiation. The radiant sensitivity, stated in microamperes/watt, is found by determining the response of the photodetector to that radiation to which the detector is most sensitive. Therefore, it is called radiant sensitivity at peak response. Sensitivity at other wavelengths can be found by using the relative response curves of Fig. 4-6. Luminous sensitivity is stated in microamperes/lumen. Here a test lamp of color temperature 2870°K is employed in place of the monochromatic source used to determine radiant sensitivity.

Conversion of lumens to watts can be accomplished only by taking into account the spectral sensitivity curve of the photodetector. These factors and other relevant parameters of photodetectors were calculated by R. W. Engstrom, some of whose data are reproduced in Table 4-1.

TABLE 4-I

Basic Data Pertaining to S-1, S-10, and S-20 Photodetectors

S-Number	Conversion Factor (lumens/watt)	Wavelength of Maximum Sensitivity	Radiant Sensitivity (μamp/W)	Peak Quantum Efficiency (percent)	Dark Current at 25° C (amp/cm^2)
S-1	94	8000	2,350	0.36	900×10^{-15}
S-10	508	4500	20,300	5.6	70×10^{-15}
S-20	428	4200	64,200	18.0	0.3×10^{-15}

4.2.1 Electrostatic Photomultipliers

One of the earliest methods used to enhance photoelectric signals was the photomultiplier. It works on the principle of secondary emission electrons produced by photoelectrons emitted from a photocathode surface.

The multiplication process increases the sensitivity by a factor of 10^6 to 10^7. This increase of sensitivity, however, does not necessarily increase the capability of the device to detect small signals, since the noise is amplified together with the signal. The detectivity, or the noise equivalent power, is the figure of merit which determines the least signal detectable by means of the photodetector.

The data of Table 4-II are found in the RCA Electron Tube Handbook concerning photomultipliers with S-1, S-10, and S-20 spectral responses:

TABLE 4-II

Noise in Photomultiplier Tubes

Tube	Spectral Response	Noise Equivalent Illumination	Temperature (°C)
6217	S-10	4.0×10^{-11} lumen	25
7162	S-1	1.7×10^{-12} watts	25
7265	S-20	7.5×10^{-13} lumen	25
7265	S-20	1.0×10^{-13} lumen	-80
7326	S-20	1.9×10^{-12} lumen	25
7326	S-20	3.0×10^{-13} lumen	-80

For tubes of S-10 and S-20 spectral response the figures of Table 4-II may be converted to noise equivalent power at the peak of the response by means of Engstrom's factors. Then the noise equivalent power for the relevant radiation may be obtained by making use of the relative response curves of Fig. 4-6. In calculating quantum efficiencies at 6943 A, for example, the frequency dependence of the quanta must be taken into account in addition to the change in response from peak to 6943 A. In this manner Table 4-III is constructed.

The information in Table 4-III may be used in answering the following question: Given the output bandwidth of 10^6 cps, what signal level is required at 6943 A to produce a s/n ratio of 1 when a 7326 photomultiplier is used at room temperature? The answer is:

$$p = (10^6)^{1/2} \times 2 \times 10^{-14} \text{ watts} = 2 \times 10^{-11} \text{ watts.}$$

This power corresponds to the incidence of

$$n = \frac{p}{hf} = \frac{2 \times 10^{-11}}{2.86 \times 10^{-19}} = 7 \times 10^7 \text{ photons/second.}$$

In 1 μs, which is the reciprocal of the bandwidth, there will be 70 photons incident. These will produce, on the average, 1.7 photoelectrons.

TABLE 4-III

Phototube Characteristics Calculated for Ruby and Neodymium Radiation*

Tube	Temperature (°C)	NEP (Peak) (1 cycle)	Relative Sensitivity for λ_o (*)	NEP λ_o W (1 cycle)	q (peak) (%)	q λ_o (%)
6217	25	7.9×10^{-14}	0.10	7.9×10^{-13}	5.6	0.36
7265	25	1.7×10^{-15}	0.22	8.0×10^{-15}	18.0	2.4
7265	-80	0.23×10^{-15}	0.22	1.0×10^{-15}	18.0	2.4
7326	25	4.4×10^{-15}	0.22	2.0×10^{-14}	18.0	2.4
7326	80	7.0×10^{-16}	0.22	3.2×10^{-15}	18.0	2.4
7102	25	1.7×10^{-12}	0.28	6.0×10^{-12}	0.36	0.04

*(λ_o = 6943 A for tubes with S-10 and S-20 response, λ_o = 1.06 μ for tubes with S-1 response.)

Sizes and weights of commercially available photomultipliers vary. One of the most sensitive is the RCA 7046 which uses 14 stages of amplification and has a 4 7/16" diamter. The required DC supply voltage is 3400 volts with an average anode current of 2.0 ma. Most photomultipliers require above 1000 volts.

4.2.2 Dynamic Crossed Field Photomultipliers

Presently available electrostatic photomultipliers are not capable of operating above a few hundred megacycles because of electron transit time spread. Recently a photomultiplication system in which energy for secondary emission multiplication is obtained from a microwave electric field has been experimentally demonstrated (*).

This photomultiplier operates on the principle of dynamic crossed field electron multiplication in which electrons are multiplied in a field configuration consisting of a high-frequency electric field and a crossed steady magnetic field; the fields are bounded by two electrodes, one of which has a high emission ratio. See Fig. 4-7.

The bandwidth of the photomultiplier is equal to one-half the electric field frequency ω due to the fact that the device is a sampling system with a sampling rate equal to the electric field frequency. Two classes of operation can be distinguished. The first in which the multiplier acts as a video amplifier from dc to $\omega/2$ and the second in which the multiplier operates as a heterodyne detector from $\omega/2$ to ω.

The gain of the photomultiplication is defined as the ratio of the current collected after n multiplication steps to cathode current. Under optimum conditions a gain of 10^8 can be obtained. The noise characteristics of the dynamic-crossed field photomultiplier should be similar to those of an electrostatic photomultiplier. Noise of the electric field source, however, may be magnified into the output.

An experimental dynamic crossed field photomultiplier has been built.* The high electric field was obtained in a resonant cavity structure. Two watts were needed at 3 gc to produce fields of 10^6 v/meter. An external electromagnet supplied a uniform 500 gauss magnetic field.

4.2.3 Traveling-Wave Phototube

There are other methods of detecting and internally amplifying a microwave modulated light beam using the principle of photoemission. Consider Fig. 4-8. Modulated light striking the photocathode surface produces a pre-bunched beam of electrons through a microwave structure. The microwave structure is designed to both amplify the modulation and remove it as an electrical signal. It may consist of a cavity, a helix or some

* O. L. Gaddy and D. F. Holshouser, Proc. IEEE, January 1963, p. 153

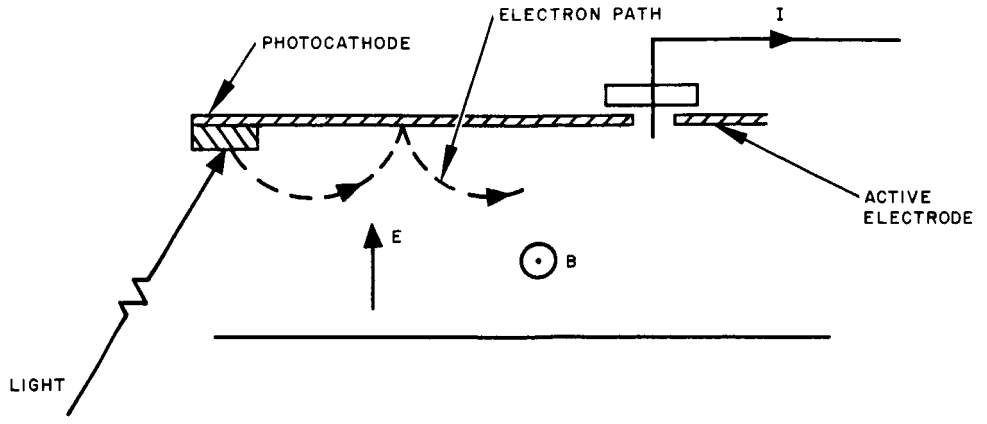

Figure 4-7. Dynamic Crossed Field Photomultiplier

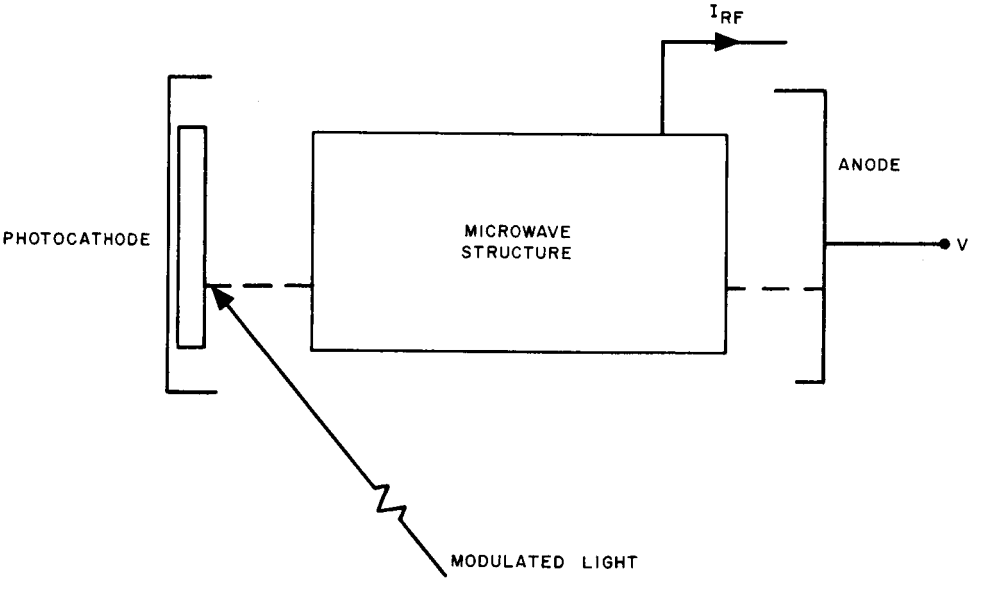

Figure 4-8. Microwave Phototube

other slow-wave circuit. A traveling-wave tube type of construction will yield a greater bandwidth capability.

Transit-time problems are not significant in these phototubes; the initial current modulation excite the well-known space charge waves, and the microwave-tube section then amplifies and detects them. The most important source of noise is shot noise.

Consider microwave amplitude modulated light striking the photocathode and causing a cathode current composed of I_1, the microwave component, and I_o the D C component resulting from the fact that the light beam isnot 100 percent modulated. If the photoelectrons are accelerated by a D C voltage, V_1.

The maximum microwave power output is given by

$$p \leq I_o \; V \tag{1}$$

The shot-noise resulting in a bandwidth B is

$$I_N^2 = 2 e \; I_o \; B \tag{2}$$

The microwave modulation signal will be

$$I_1^2 = \left(\frac{\eta n e}{\sqrt{2}}\right)^2 \tag{3}$$

where n is the number of arriving photons to the cathode and η is its quantum efficiency.

Equating (2) and (3) the minimum detectable number of photons can be found. Thus the minimum detectable power is

$$P_1 = nhf = \frac{2 h c}{\eta \lambda} \left(\frac{I_o B}{e}\right)^{1/2} = \frac{(I_o B)^{1/2}}{\eta \lambda} \times 10^{-15} \text{ watts} \tag{4}$$

A plot of equation (4) is shown in Fig. 4-9. Minimum detectable power is plotted as a function of modulation bandwidth frequency and direct current in the phototube. The wavelength is taken at 0.5 microns and $\eta = 0.1$ for a typical photocathode. The minimum detectable power as may be seen, is quite high with d. c. current present. An improvement in NEP can be obtained by superheterodyne as will be noted later.

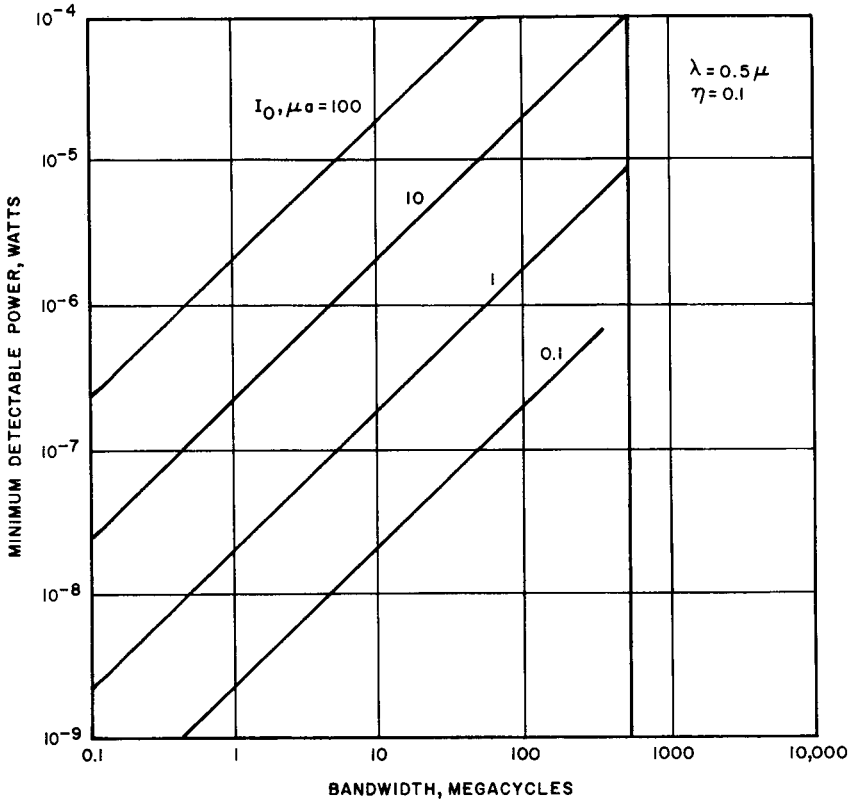

Figure 4-9. Minimum Detectable Power of a Microwave Phototube as a Function of Modulation Bandwidth and dc Cathode Current

(x = 0.5μ, quantum efficiency 0.1)

Sylvania has recently been marketing traveling wave photo tubes. They are all permanent magnet types, light weight and compact. Weight including integral magnets is below 5 lbs. and the sizes are approximately 2" diameter x 15" length. The types of photocathode available are either the S-1 or conventional thermionic cathode. Phototubes of this type are available spanning a frequency range between 1 and 12.5 gc. small signal gain between 20 and 30 db should be possible. For further details refer to Table 4-IV.

TABLE 4-IV MICROWAVE PHOTOTUBES

Type	Frequency (gc)	Helix Voltage V DC	Thermionic Cathode Current μA DC
SYD - 4301	1 - 2	250	400
SYD - 4302	2 - 4	450	400
SYD - 4303	4 - 8	-	-
SYD - 4304	4 - 12.5	1030	400

4.3 COMPARISON OF PHOTODETECTING COMPONENTS

A comparison table of the various types of photodetectors discussed in this section is given in Table 4-V. Photomultipliers offer the highest sensitivity of all detectors. Traveling wave phototubes, although capable of high frequency response and large bandwidths, are limited in sensitivity by shot noise when used as envelope detectors.

TABLE 4-V

COMPARISON OF DIFFERENT TYPES OF PHOTODETECTORS

	Advantages	Disadvantages
Photo Diodes	1. Has high sensitivity NEP 10^{-9} watts typically at 5 mc 2. High quantum efficiency (from 0.5 to 1.0) which extends into the IR 3. Low dark current 4. Simple and rugged 5. Video bandwidth capability 6. gc frequency response capability	1. For high frequency requires cavity resonator. 5-10 percent bandwidth 2. Does not have internal amplification (except if used as parametric amplifier). Shot noise amplified in external amplifying system 3. Small detection area. Transit time imposes tradeoff between sensitive area and diode capacitance 4. Low equivalent resistance at high frequencies (50 ohms).
Photomultipliers	1. Has high sensitivity NEP 10^{-11} watts at 5 mc 2. Internal amplification by electron multiplication 3. gc frequency response capability by dynamic crossed field multiplication 4. Broad bandwidth capability	1. Low quantum efficiency 2. Requires cooling for low dark current 3. Infrared response limited to 1.1 microns 4. When used for optical superheterodyne reception direct current from L.O. is amplified causing burn out 5. Not rugged. Mechanically complex 6. Requires high anode voltages (>1 kv)
Microwave Phototubes	1. Internal amplification by traveling wave interaction 2. When used for optical superheterodyne direct current aids amplification 3. Ruggedness comparable to microwave tubes 4. gc frequency response capability 5. Broad bandwidth capability (3:1) 6. No transit time effects 7. High equivalent resistance (100K)	1. Low quantum efficiency 2. Limited infrared response (1.1 microns) 3. Requires high anode voltages 4. Not suitable for video frequency detection

5.0 DEMODULATION SYSTEMS

The optical receiver may make use of several different detection systems which have already been well defined for RF receivers. The techniques are the same, only the components are different. The actual components to be used have been described in detail in Section 4. Therefore, this section will be concerned with their combination into systems to detect information on an optical carrier. Three systems will be considered - direct video detection, optical heterodyne and optical homodyne.

5.1 VIDEO DETECTION

Consider the operation of a direct detection receiver as shown in Fig. 5-1. The incoming signal of power P_s passes through an optical filter and impinges on a photosensitive surface. If η is the quantum efficiency, e the charge on the electron, then the current which flows, I_1, is

$$I_1 = \eta e \frac{P_s}{hf} \tag{1}$$

The power of the generated signal is proportional to the square of the current. If the incoming signal is amplitude modulated, the resulting current is amplitude modulated. This current may or may not be amplified, depending on the particular detection device. The amplitude modulation is picked off directly as an electrical signal. This is one of the simplest systems to implement but has inherent limitations. Filtering must take place at optical frequencies and good optical filters of a bandwidth less than 5A (angstrom) are not available. At λ = 1 micron 5A bandwidth corresponds to a frequency bandwidth of greater than 150 gc. It is usually desired to pick out weak optical signals against a fairly strong sunlight or other noise background. A direct-detection receiver then must operate with a wide predetection bandwidth, even through the signal bandwidth is relatively narrow.

There will be three important noise contributions in the video detector. The quantum noise associated with the detection of the modulated laser beam, the thermal or Nyuist noise of the load resistance, and the shot noise which depends on the dark current. It is assumed that there is no background noise. Assuming the noise currents are uncorrelated and 100 percent modulation of the light beam

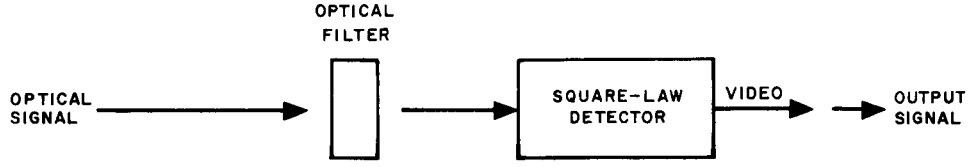

Figure 5-1. Video Detection Receiver

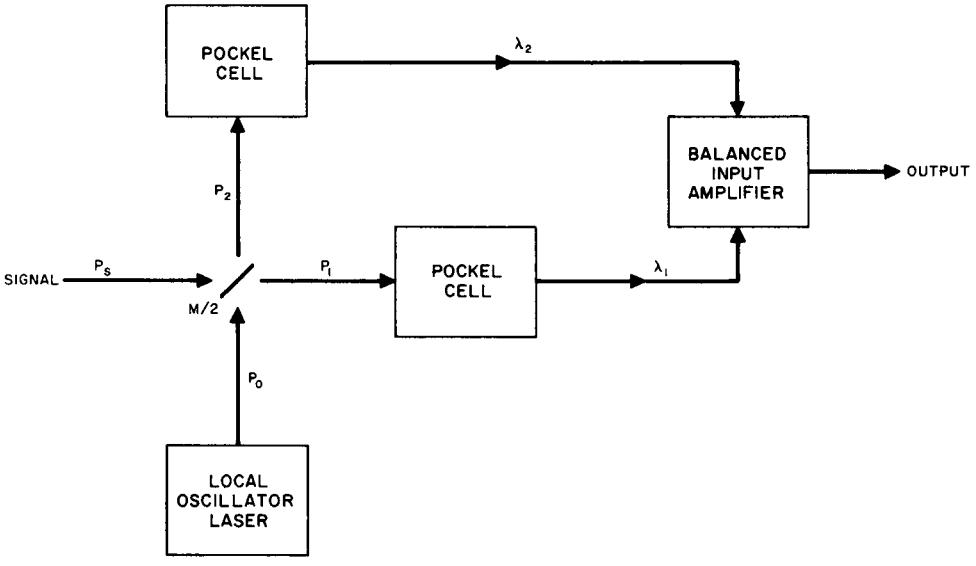

Figure 5-2. Balanced Photoelectric Mixer

$$\left(\frac{S}{N}\right)^2 = \frac{(\eta e\, n)^2}{I_N^2 + I_R^2 + I_D^2} \tag{2}$$

Where I_N, I_R, I_D are the quantum, thermal and shot noise of the device, the quantum noise current is

$$I_N^2 = 2\, e^2\ \eta\, n\, B \tag{3}$$

assuming a Poisson distribution for received photons. The thermal noise is equal to

$$I_R^2 = \frac{4\, K\, T\, B}{R}$$

and the shot noise is

$$I_D^2 = 2\, e\, I_o\, B$$

where I_o is the dark current. The 1/f noise will influence the low-frequency region of the modulation and should also be included. If the dark current and I_R is sufficiently small compared with I_N, it can be shown that the number of incident photons required for unit signal to noise ratio is

$$n = \frac{2\, B}{\eta}$$

5.2 OPTICAL SUPERHETERODYNE AND HOMODYNE

Optical superheterodyne and homodyne systems perform the same functions as their RF counterparts, namely conversion to some IF frequency at which filtering and detection take place. Fig. 5-2 shows the block diagram of a balanced heterodyne, using a local oscillator laser of power P_o. A half-silvered mirror is used to superimpose the signal and local oscillator beams before they hit a photocathode surface.

In Fig. 5-2, the half-silvered mirror, M/2, acts as a hybrid junction and has reflecting and transmitting coefficients of

$$e^{j\pi/4}/\sqrt{2} \qquad \text{and } e^{-j\pi/4}/\sqrt{2},$$

respectively. If

$$P_s = \left[A_s \cos (\omega_s t + \pi/4) \right]^2$$

$$P_o = \left[A_o \cos (\omega_o t - \pi/4) \right]^2 \qquad (4)$$

then from Olivers* analysis

$$P_1 = \frac{1}{2} (A_o \cos \omega_o t + A_s \cos \omega_s t)^2$$

$$P_2 = \frac{1}{2} (A_o \sin \omega_o t - A_s \sin \omega_s t)^2 \qquad (5)$$

The signal output current i_s is given by

$$i_s = i_1 - i_2 = \frac{\eta e}{2hf} A_o A_s \cos \delta t \quad \text{where} \quad \delta = \omega_s - \omega_o \qquad (6)$$

i_s has the mean square value

$$\overline{i_s^2} = \left(\frac{\eta e}{hf} \right)^2 \frac{A_o^2 A_s^2}{2} \qquad (7)$$

If the bandwidth of the circuit following the mixer has a bandwidth B, the mean square shot noise may be written

$$\overline{A_N^2} = 2eB (i_1 + i_2) = \frac{\eta e^2 B}{hf} (A_o^2 + A_s^2) \qquad (8)$$

Taking the ratio of (6) to (7)

$$\frac{S}{N} = \frac{\eta}{2hfB} \frac{A_o^2 A_s^2}{A_o^2 + A_s^2} \qquad (9)$$

As A_o is increased, both the S/N ratio and signal power increase, the second without limit, the first approaching

$$\left. \frac{S}{N} \right|_{max} = \frac{\eta}{hfB} \frac{A_s^2}{2} = \frac{\eta \overline{P_s}}{hfB} \qquad (10)$$

* "Signal to Noise Ratios in Photoelectric Mixing", B.M. Oliver, Proceedings IRE, December 1961, p. 1960.

where $\overline{P_s}$ is the average of P_s. The number of photons required for unity signal to noise ratio is thus $n = B/\eta$.

If $hf \gg KT$, then the S/N ratio of an ideal amplifier takes the form

$$\frac{S}{N} = \frac{\overline{P}}{h f B} \tag{11}$$

If $\eta = 1$, (10) is simply the S/N ratio of an ideal amplifier.

The above results apply when $\omega_o \neq \omega_s$ (a superheterodyne receiver). If we consider a homodyne receiver in which $\omega_s = \omega_o$ and $\delta = 0$, from (6) $\overline{i_s^2}$ is twice as great as before, $\cos \delta t$ having been replaced by unity. Equation (8) is still true so the S/N ratio has doubled:

$$\frac{S}{N}_{max} = \eta \frac{\overline{P_s}}{\frac{h}{2} f B}$$

With a quantum efficiency of unity, the signal to noise ratio is now twice as great as an ideal amplifier. This improvement in S/N ratio with homodyne mixing as compared to heterodyne mixing arises because the signal coversion gain is doubled while the noise level, which occurs after mixing, is not affected. For unity signal to noise ratio the number of photons required is $n = B/2\eta$. This is not true when background noise is received with the signal. Any change in gain effects both signal and noise equally.

If heterodyne action is to be obtained the signal and local oscillator wavefronts must be parallel and in phase at the cathode. If this were not the case the resulting intermediate frequency signals would have different relative phases and would tend to cancel out. The beams must be parallel to within

$$\Delta\theta = \frac{\text{wavelength}}{\text{Photocathode diameter}}$$

This results in a very small angle (for $\lambda = 0.5\mu$ and the photocathode diameter equal to 1 cm, $\Delta\theta$ is 10 seconds of arc). Precise optical alignment is then required increasing the complexity of the system.

A number of important properties of the heterodyne and homodyne receivers are apparent from the above analysis. In the first place the signal to noise ratio is independent of the local oscillator strength and of the size of the dark current, an important consideration in video detectors. Thus without degrading the signal to noise ratio large DC currents may be allowed thus aiding in the amplification process in the case of the photoemissive devices discussed in Section 4-2.

Heterodyne reception increases the sensitivity of the detector by at least a factor of two and by a factor of four for homodyne reception over video detection. It also permits frequency filtering at the intermediate frequency where methods are well established rather than at the optical frequency. This is an important consideration when operating in the presence of background noise. A 1 gc bandwidth, for instance, corresponds to about 0.016 A at 0.7 microns wavelength. Good optical filters below 5 A are not presently available.

However these advantages should be weighed against the added size and complexity of the system as compared to video detection. The frequency of the local oscillator laser must be held fairly stable otherwise the optical beat frequency will drift. Fig. 5-3 indicates a type of electrical circuitry for use with the optical heterodyne where the voltage controlled laser local oscillator is tuned in frequency by adjustment of its reflectors so that the radio frequency beat is maintained constant.

One method which has been suggested* to stabilize angular misalignment between the local oscillator laser and incoming waves is shown in Fig. 5-4. The pinhole and lens assembly is aligned to focus the signal beam at the pinhole. The pinhole then selects precisely that local oscillator wave which is specially coherent with the incoming signal wave over the photodetector surface.

Heterodyne reception is one of the practical methods used presently to detect frequency or phase modulated light. The local oscillator provides the frequency reference with which to compare the modulated wave. The frequency modulated beat frequency is demodulated by conventional FM receiver techniques.

A technique for converting phase modulated beams to amplitude modulation has been demonstrated using an optical homodyne system**. With proper adjustment of the phase between the local oscillator laser and the signal beam the PM sidebands become AM sidebands of the local oscillator. Detection of the latter modulation is conventional. Homodyne amplification results from the fact that the resulting photocurrent is proportional to the square root of the local oscillator beam intensity. A very stable local oscillator frequency is, of course, required in this system. Another disadvantage is that a large phase modulation index results in a sinusoidal rather than a linear representation of the phase modulation.

---

* "A Laser Design for Space Communications", L. Goldmuntz, IRE Convention Record, September 1962, p. 305.

** P. Rabinowitz, et. al., Proceedings IRE, November 1962, p. 2365.

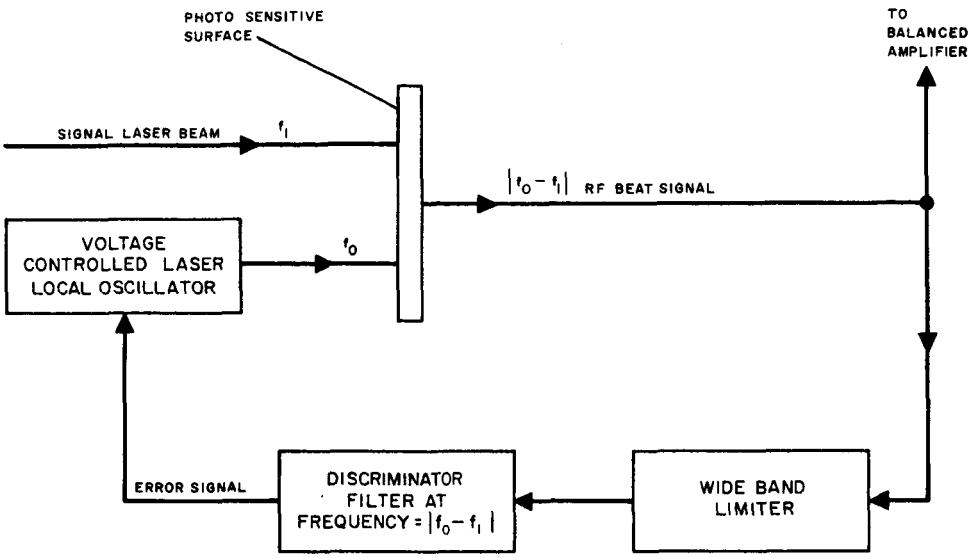

Figure 5-3. Heterodyne Optical Receiver with Voltage Controlled Local Oscilaator Laser

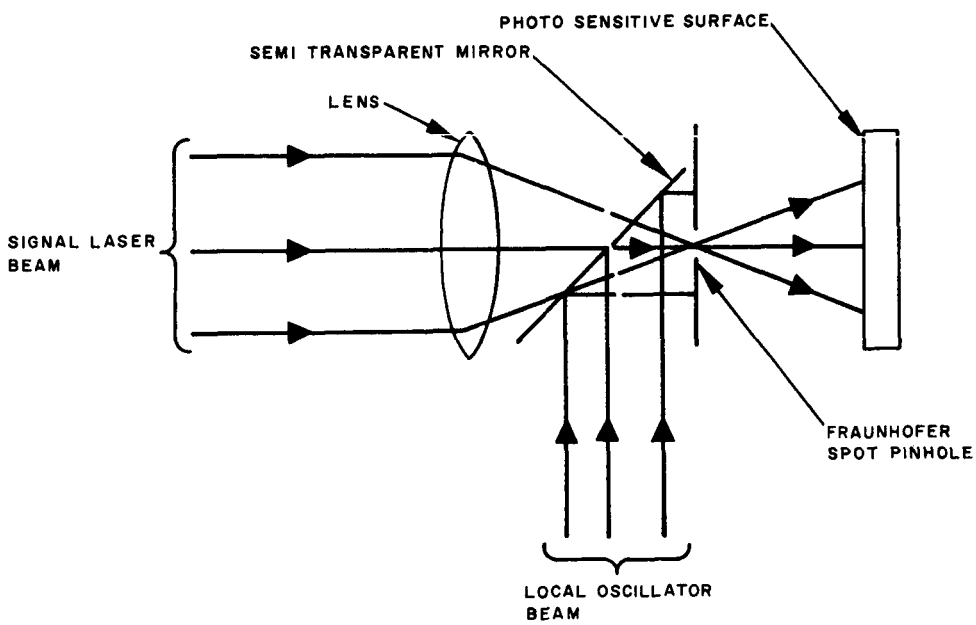

Figure 5-4. Method to Stabilize Misalignment in Optical Heterodyne Receiver

5.3 ELECTRO-OPTICAL HETERODYNE

The frequency of amplitude or phase modulation on a light beam can be shifted to a new value by passing the beam through an electro-optic modulator.

Fig. 5-5 is a schematic of a heterodyne demodulator for an intensity modulated light beam. An intensity modulated signal of a frequency f_s is mixed by an electro-optical light modulator of the type described in Section 2.1 driven at a frequency f_o. It can be shown* that the light beam has a frequency component at $| f_o - f_s |$ of amplitude equal to

$$A = M \cos \frac{\pi V_B}{V} \quad J_2 \left(\frac{\pi V_o}{V}\right)$$

where M is the modulation index of the intensity modulated signal, V_B is the DC bias voltage across the crystal, V_o is the amplitude of the local oscillator voltage and V is the voltage necessary to produce a half wave retardation. The minimum conversion loss is 3.44 (approximately 5.4 db).

In the case of polarized modulated light the minimum conversion loss is only 2.4 db.

Electro-optical heterodyning provides a convenient method of shifting microwave modulated light to video frequencies thus permitting the use of photo-detectors which only operate within this frequency range. The same signal to noise relations discussed in Section 5.1 for video detection would therefore be the case here.

5.4 DEMODULATION OF POLARIZATION MODULATION

As discussed in Section 1, polarization modulation offers the advantage over intensity modulation that there is no loss of the laser signal in the process of modulation and demodulation. Fig. 5-6 shows a polarization modulation receiver. The Wollaston prism acts as though two crossed polarization analyzers were acting simultaneously splitting the beam I_o into two mutually polarized beams I_1 and I_2. Their intensity is of the form

$$\frac{I_1}{I_o} = \frac{1}{2} + \sin \delta$$

$$\frac{I_2}{I_o} = \frac{1}{2} - \sin \delta$$

where δ is the optical phase retardation

---

* D. J. Blattner and F. Sterzer, RCA Review, September 1962, p. 407.

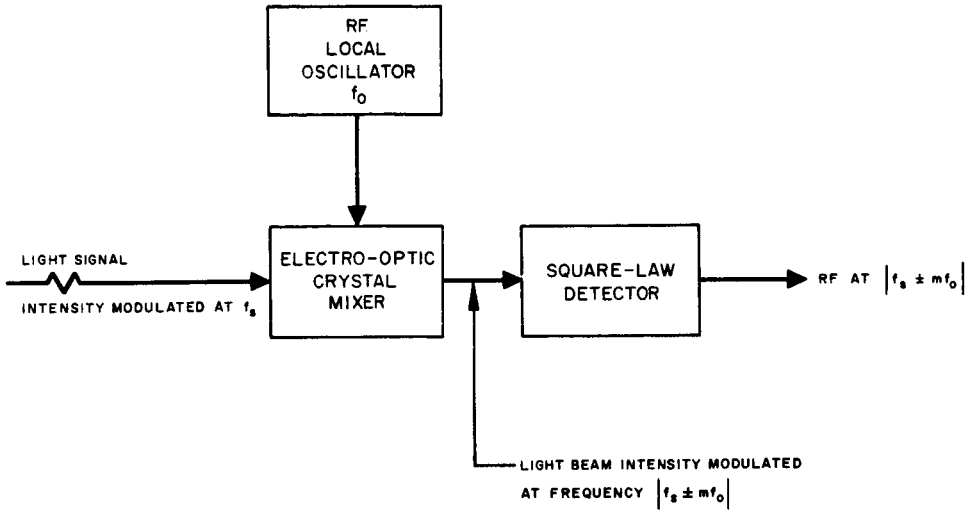

Figure 5-5. Heterodyne Light Receiver for Narrow-Band Detectors

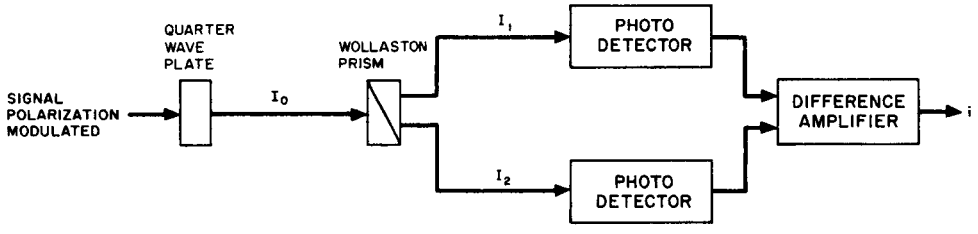

Figure 5-6. Polarization Modulation Receiver

$$\delta = \frac{\pi}{\lambda} \quad n^3 \; r \; V \sin \omega t$$

λ being the light beam's wavelength, n the index of refraction of the Pockels crystal, its electro-optic constant and V sin ω t the electric modulation impressed on the crystal. The signal current from the photodetectors will be proportional to I_1 and I_2 the light intensity striking their surface. The output of the photodetectors is fed into a difference amplifier and the resulting current is then

$$i = K \, 2 \, I_o \sin \delta$$

where K is a constant which takes into account the photo electric conversion and the amplifier gain.

Background noise is reduced in this system since over extensive periods of time the average noise levels cancel. If a fraction of the current from the photodetectors is fed into a summing amplifier, the signal output would be of the form

$$i = K' \, I_o$$

this would allow in the absence of background noise an estimation of the stability of the transmitting laser source.

5.5 COMPARISON OF DEMODULATION SYSTEMS

Table 5-I offers a summary of the most important advantages and disadvantages of the demodulation systems discussed in this section. Although heterodyne and homodyne reception has a number of advantages over envelope detection present state of the art is not yet sufficiently advanced to make these systems practical. It would therefore appear that for near future applications envelope detection offers the best possibilities.

TABLE 5-1

TYPES OF DEMODULATION SYSTEMS

	Advantages	Disadvantages
Video Detection	1. Simple straightforward system. 2. Subcarrier demodulation is possible.	1. Filtering takes place at optical frequencies. 2. Has lowest signal to noise ratio as compared to other systems. 3. Cannot be used to detect F. M.
Superheterodyne	1. Permits filtering at I. F. frequency. 2. Improved signal to noise ratio specially in presence of background noise. 3. Angular discrimination. Very narrow beam receiver. 4. Heterodyne amplification takes place since signal power is proportional to L. 0. intensity. 5. Dark current is suppressed. 6. Suitable for F. M. detection.	1. Increased system complexity. 2. Requires stable local oscillator and alignment between L. 0. and signal.
Homodyne	1. Optimum signal to noise ratio. 2. Homodyne conversion gain. 3. Suitable for F. M. detection.	1. Same disadvantages as superheterodyne. 2. Very stable local oscillator is required.
Electro-optical Heterodyne	1. Provides optical method of shifting microwave frequency to video frequency. 2. Eliminates subcarrier before demodulation. 3. Allows use of conventional video frequency photodetector.	1. Increased complexity of system over video detection. 2. Conversion loss. 3. Signal to noise ratio is decreased as compared to video detection.

6.0 TWO TYPICAL TRANSMITTER-RECEIVER SYSTEMS

This section will be concerned with two types of laser systems - a pulsed laser and a cw laser source. We will assume the following parameters:

cw laser

low power ($\sim$ 1 watt)

pulsed laser

high power ($\sim$ 10 kw)

1 pulse per second

pulse duration = 1 milisec.

6.1 C-W LASER

Since the power output of the c-w laser is limited, the modulating bandwidth must be as narrow as possible. Existing slow scan video cameras can transmit on less than a 5 kc bandwidth. Total bandwidth chosen, including 10 telemetry and 1 voice plus guard bands is 8 kc at a 100 kc center frequency. The individual channels are frequency multiplexed and summed before being fed into the modulator. An estimated 5 watts is required to produce a 10 mw input into the modulator amplifier.

Polarization modulation was chosen as the modulation system because of low modulator insertion loss. 6 watts at 600 v is necessary to drive a 3 stage Pockels cell modulator at 100% modulation. This system is shown in Figure 6-1.

The polarization modulation is detected by a Wollaston prism and photomultipliers. By using comparison amplifiers, the effects of optical carrier variation can be eliminated.

The subcarriers are separated by band pass filters and demodulated. Low pass filters eliminate high frequency components. The receiver system is shown in Figure 6-2.

6.2 PULSED LASER

Since the pulsed laser can transmit for only a short time, information storage facilities are necessary. For an 8 kc bandwidth, 1 m sec pulses at 1 pps, an 8 mc readout is required. 10 mc might be used to allow reset time.

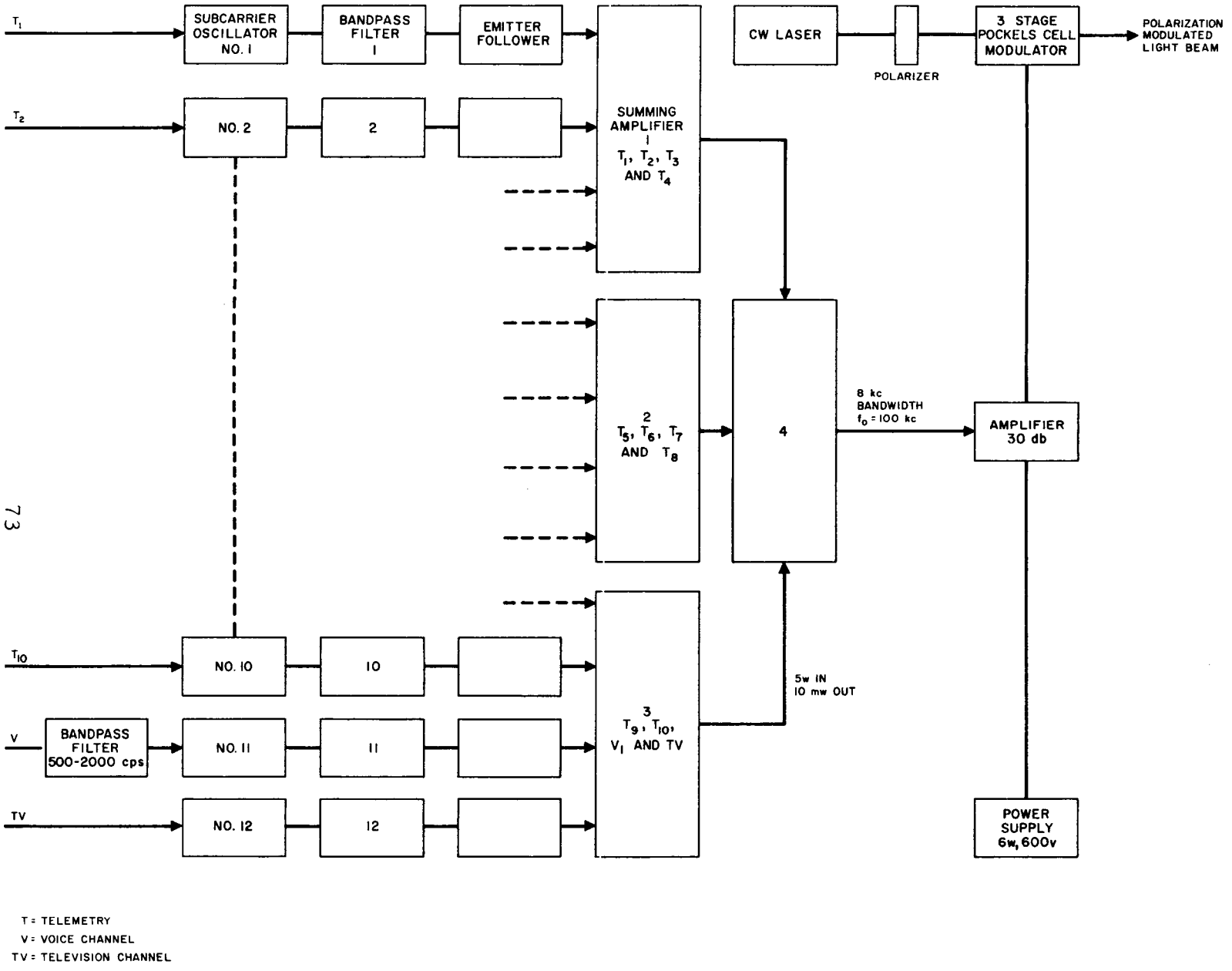

Figure 6-1. CW Laser Transmitter

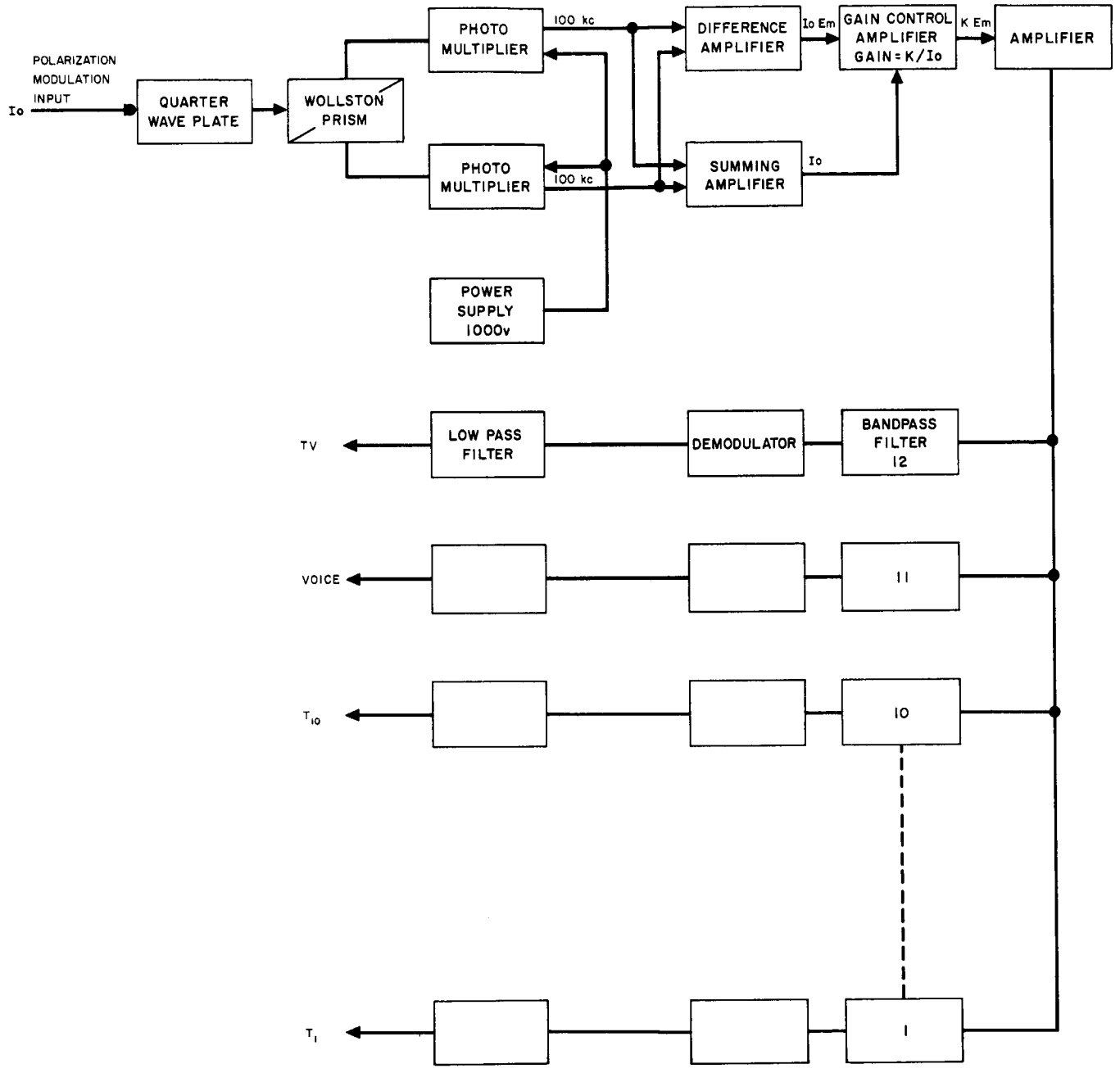

Figure 6-2. CW Laser Receiver

The light beam is intensity modulated at a 3 gc subcarrier frequency. A microwave subcarrier frequency is chosen to reduce modulator power requirements. The traveling wave modulator would require 10 w for 30% modulation.. This system is shown in Figure 6-3.

A traveling wave phototube is used to detect the intensity modulated light signal. The 10 mc information is retrieved by conventional superheterodyne techniques. The modulated light beam will arrive at the receiver in short bursts, necessitating storage to provide a smooth information rate.

A pulsed laser receiver system is shown inFigure 6-4.

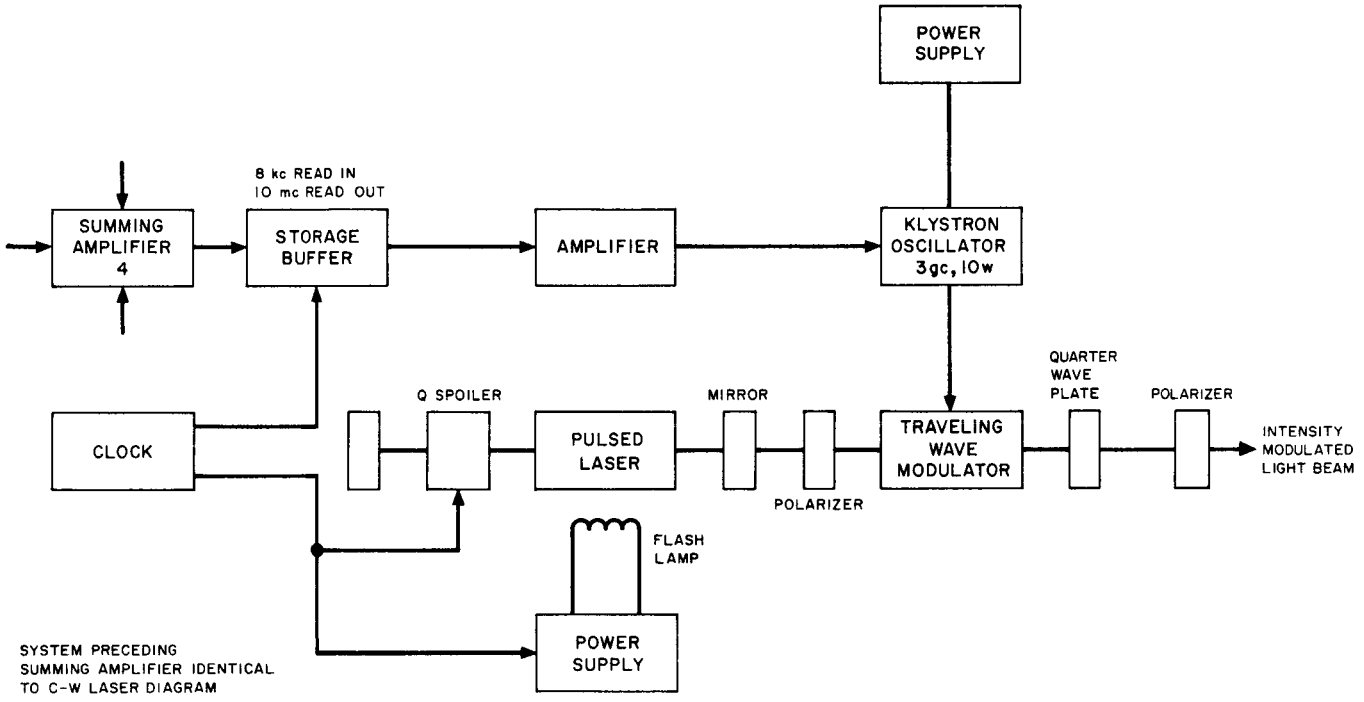

Figure 6-3. Pulsed Laser Transmitter

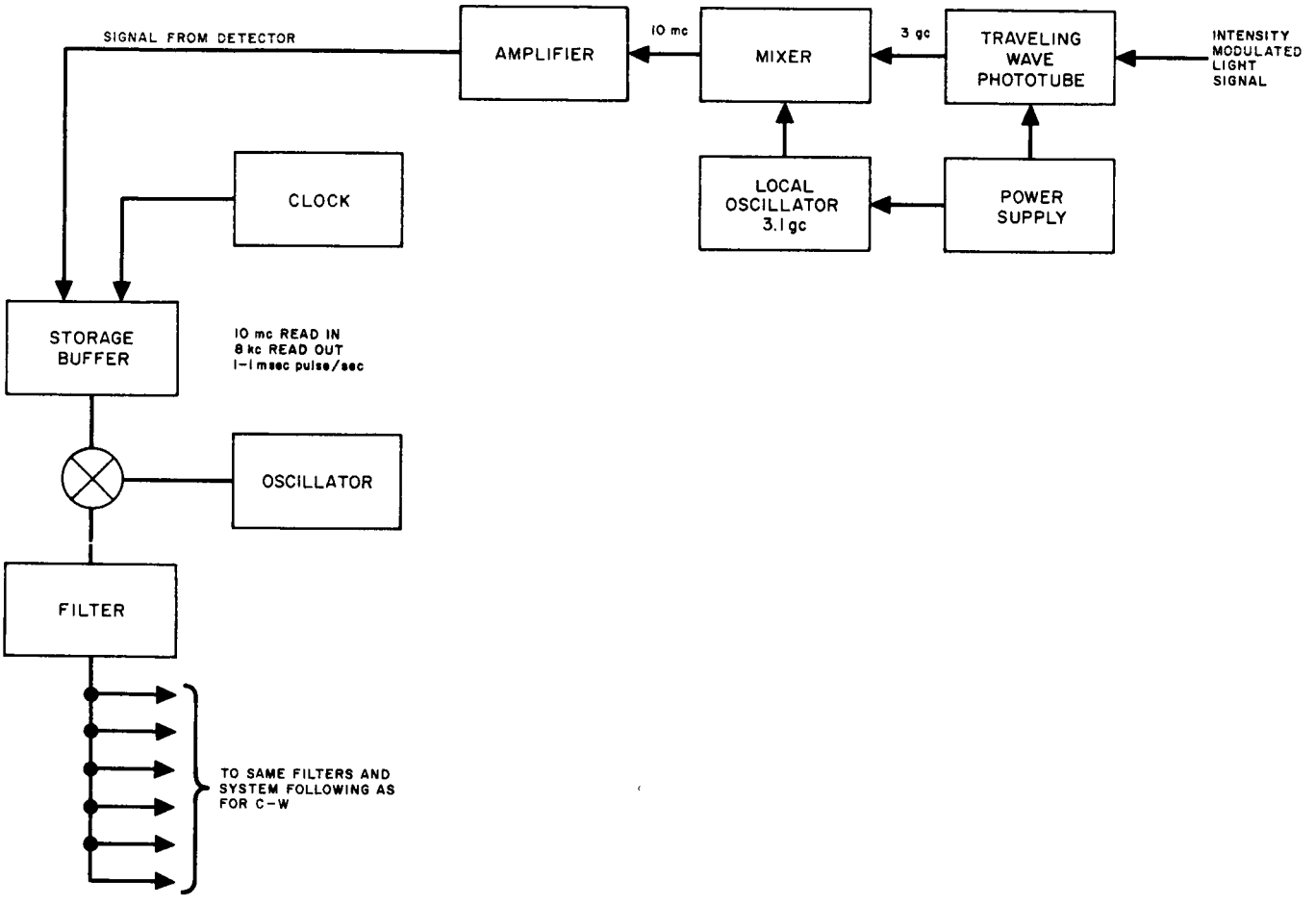

Figure 6-4. Pulsed Laser Receiver

7.0 SYSTEM CONSIDERATIONS

This section is a continuation of the research into basic areas which serve to establish guidelines for the system evaluation which follows. In particular, atmospheric attenuation is discussed and the salient features delineated. Then some rather critical observations are made in regard to the tracking and pointing problem and its relevance in achieving a workable system

7.1 ATMOSPHERIC ATTENUATION IN VISIBLE & INFRARED

The attenuation of a parallel beam in the visible and near infrared frequencies can be attributed to two causes, scattering and absorption. The effect of each of these will be discussed separately. The frequency region being considered will be from a wavelength of 0.4 μ to 14 μ (1 $\mu = 10^{-6}$ meters $= 10^4$ angstrom units).

7.1.1 Analysis of Attenuation Curves

The solid attenuation curve shown in Fig. 7.1 is constructed after Gebbie's data [3]. This is one of the best known and most thorough investigations of infrared transmission. Gebbie's data was taken over two transmission paths of 2264 yds. and 4478 yds. respectively. Three frequencies were chosen as reference and long time measurements were made at these frequencies. Then when scintillations were at a minimum, frequency sweep measurements were made. Using the three frequencies as reference, a plot of transmission per nautical mile (2000 yd.) was constructed. In Fig. 7.1 this has been converted to attenuation in db per mile. Gebbie assumed that his bandwidth was narrow enough so that the exponential law held in relating his measurements over a 4476 yd. path to transmission per nautical mile. He used a spectrometer receiver, and two different crystal prisms, resulting, respectively, in a receiver bandwidth of 0.02 μ (200 Å), and 0.1 μ (1000 Å). Although the attenuation curve looks smooth in the high absorption regions, higher resolution shows each band to be made up of many small, almost equally spaced, absorption lines. To determine effects of each line would require a bandwidth of 1 angstrom or less [7]. It seems doubtful that Gebbie's data, taken in the high absorption regions, can be extrapolated to path lengths much greater or much less than the experimental path. The curve in Fig. 7.1 only gives an idea of the attenuation to be expected in these regions.

The curves of Fig. 7.2 were extracted from Gebbie's data but were not presented by Gebbie himself. Gebbie took data relating transmission at infrared and transmission at 0.61 μ, in the visible. From this data, the expression

$$T = e^{-\sigma x}$$

where

$$\sigma = k \lambda^n$$

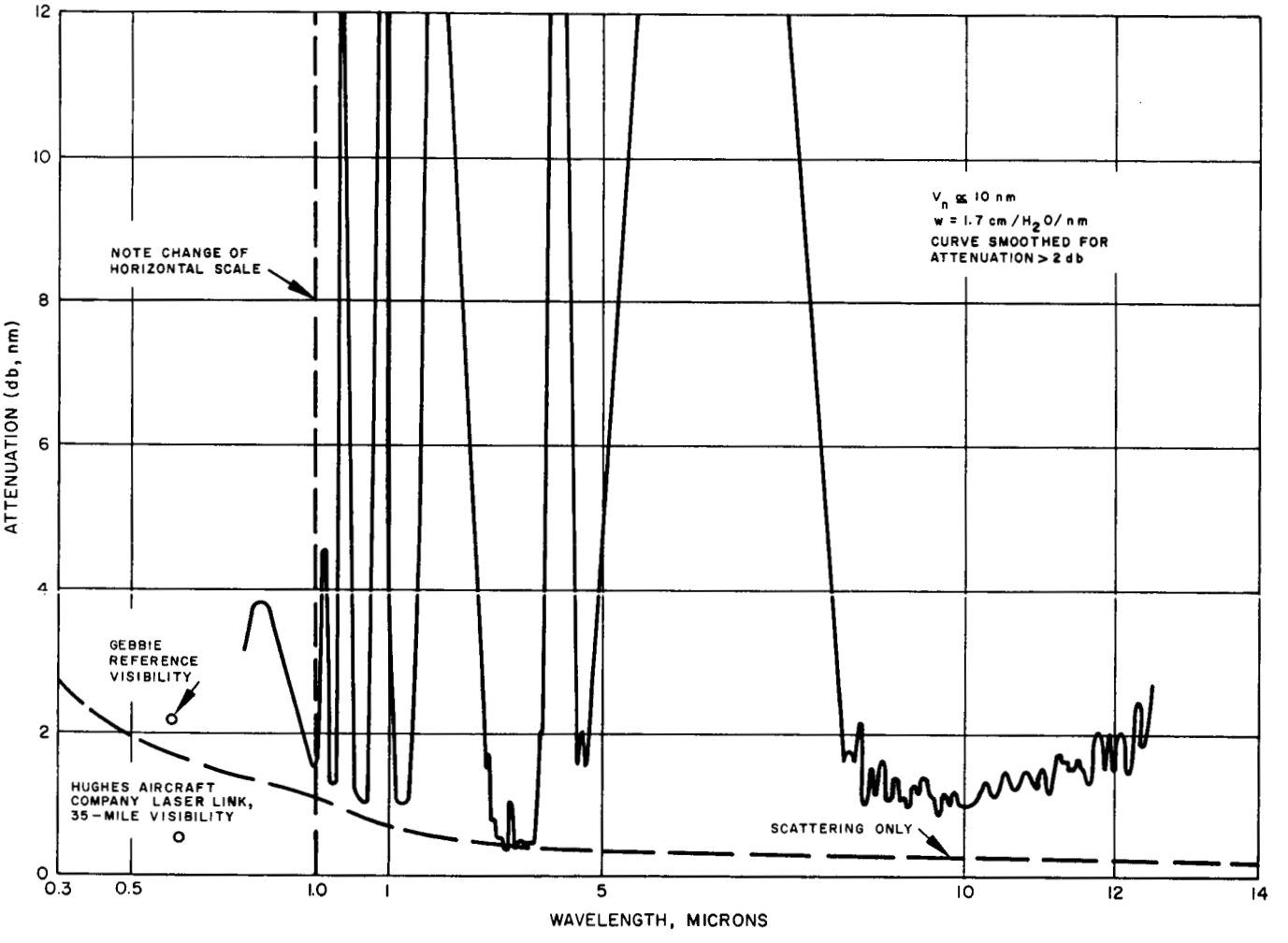

Figure 7-1. Total Atmospheric Attenuation

fits to a very good approximation when n = -0.7. k is a constant dependent on weather conditions and x is the path length. Then if we consider the transmission at 0.61μ,

$$e^{-k\ \lambda^{-0.7}} = T$$

$$\ln T = k\,\lambda^{-0.7}$$

$$k - \lambda^{0.7}\ \ln\frac{1}{T} = 0.707\ \ln\frac{1}{T}$$

and k can be found as a function of transmission at 0.61μ. Then for various values of transmission, which is expressed here in terms of visibility, curves can be plotted of attenuation vs. wavelength. These are the curves shown in Fig. 7.2.

Now the problem is to relate visibility and transmission. The contrast for a black object of luminance B_b seen against a horizon background of luminence B_n is

$$C = \frac{B_n - B_b}{B_n}\quad e^{-Bx}$$

where B is the total attenuation coefficient. If the contrast is equal to the threshold value for the human eye, the distance x is the visual range. If C = 0.02, the generally accepted threshold for the human eye, then substituting above

$$V_n = 3.912/B$$

This threshold value of C varies with background luminance, increasing during twilight and night. It also depends on the angular diameter of the object, which must be greater than 1/2 minute of arc for these formulas to apply.

Any attempt to use these charts to determine absolute attenuation values must take into account the very subjective nature of the parameters which must be used, particularly in regard to the Mie scattering attenuation. However, the figures do indicate trends and will give reasonably accurate numbers. The attenuation found for the HAC Communications laser link from Baldwin Hills to the Malibu Research Labs shows reasonable agreement with Fig. 2. The accuracy with which these curves may be used will usually be limited by the accuracy with which the meteorological parameters can be determined. Since atmospheric density and pressure have an effect on both scattering and absorption, these curves are valid only where the data was taken, in a horizontal path at sea level.

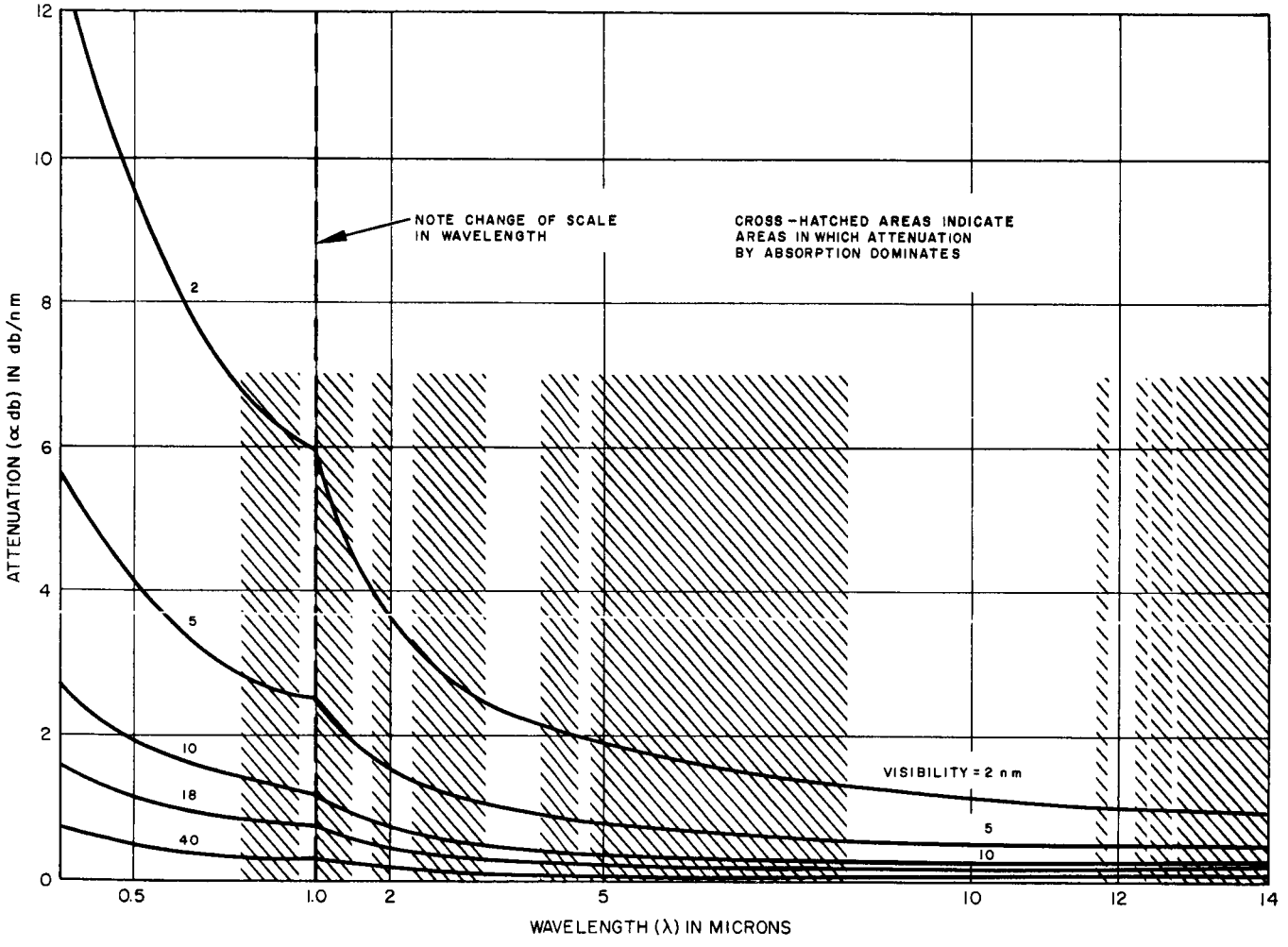

Figure 7-2. Atmospheric Attenuation due to Scattering versus Wavelength

7.1.2 Atmospheric Absorption

The transmitter power which gets to the receiver may be expressed as a function of x, the path length, and λ, the wavelength as

$$d^2I(x,\lambda) = -\alpha(\lambda)\, I(x,\lambda)\, dx\, d\lambda$$

where I is the intensity at a given x and λ, and α is an attenuation factor which is dependent on wavelength. Integrating with respect to x

$$dI(R,\lambda) = I_o(\lambda)\, e^{-\alpha(\lambda) R}\, d\lambda$$

$I_o(\lambda)$ appears as a constant of integration and represents the initial intensity which may also be a function of wavelength. Integrating once more

$$I(R,\lambda) = \int_o^\infty I_o(\lambda)\, e^{-\alpha(\lambda) R}\, d\lambda$$

The transmission factor for a beam is defined as

$$T = \frac{I(R)}{I(O)} = \frac{\int_o^\infty I_o(\lambda)\, e^{-\alpha(\lambda) R}\, d\lambda}{\int_o^\infty I_o(\lambda)\, d\lambda}$$

If we assume that α, the attenuation factor is not a function of wavelength then

$$T = e^{-\alpha R}.$$

This is the function which is usually used to represent the percent of a transmitted beam which is received at a distance R. If $I_o(\lambda)$ defines the limits of integration to be very close together, and α is relatively constant over the range of integration, the exponential law holds. This is true of monochromatic beams or when attenuation is due solely to scattering.

It is evident that if $\alpha(\lambda)$ varies rapidly in the region of integration, then the transmission is not going to follow the exponential law. This is the case when atmospheric absorption is measured with broadband sources. For

this reason much of the experimentally determined data previously obtained in regions of high atmospheric absorption (due mostly to water vapor and CO_2) cannot readily yield the value of the absorption coefficient α. Thus experiments utilizing other than narrowband sources do not lead to a prediction of the transmission losses expressed in dbs per mile. With the advent of lasers such high intensity narrow band sources are now available. In the absorption bands, the value of α for a given λ depends linearly on the density of H_2O and CO_2 in the beam path. Visibility, on the other hand, is essentially independent of these densities, except for relative humidity greater than 85%, and depends mostly on Mie scattering. Thus several parameters, such as temperature, relative humidity and path length are required to determine absorption effects. However, attenuation is usually very high in the absorption regions and these areas will probably be avoided for transmission through the atmosphere.

Nevertheless, the possibility of narrow transmission windows within the broad absorption regions should not be entirely discounted, particularly at higher altitudes where molecular line broadening due to collisions should sharply decrease.

7.1.3 Atmospheric Scattering

Attenuation in the atmosphere due to scattering may be divided into two types, Rayleigh scattering and Mie scattering. Rayleigh scattering is due to particles which are small compared to the wavelength. Mie scattering is due to particles comparable in size to the wavelength. The radii of various air particles are listed in Table 1.

Table 1: RADIUS OF VARIOUS AIR PARTICLES

Smoke (general)	0.001 - 0.1 μ
Smoke (from burning oil)	up to 1.0 μ
Haze (general)	0.001 - 0.1 μ
Fumes (general)	0.1 - 1.0 μ
Fumes (industrial smelter)	100 μ
Dust (general)	1 - 10 μ
Fog	5 - 50

Rayleigh scattering particles follow a relatively constant and predictable distribution in the atmosphere so that for a given altitude and distance the

transmission may be written as

$$T = e^{-\beta x}$$

where β is a function of altitude and frequency. This same expression is valid for Mie scattering but the attenuation factor is no longer predictable. The Mie atmosphere depends on many factors which do not lend themselves to analytical expression. However, the accompanying graphs present one method of obtaining some indication of the magnitude of the scattering coefficient based on experiment, and hence the total attenuation per unit distance.

7.1.4 Vertical Paths and Optical Thickness

The previous discussion has been concerned with attenuation along horizontal paths. For deep space communications, attenuation along paths near zenith will be of particular interest.

Consider a monochromatic beam of light originating above the atmosphere and directed vertically downward. Let this beam strike the "ground" at altitude ℓ_o above mean sea level; here and in the rest of the paper we use the word "ground" to mean any surface which sets a lower limit to the transparent atmosphere, whether the surface is earth or water or the top of a cloud bank. If $\gamma = e^{-\tau(\ell_o)} = e^{-\tau_o}$ is the fraction of the energy in this vertical monochromatic beam which reaches the ground unscattered, then the dimensionless number $\tau(\ell_o) = \tau_o$ is called the optical thickness of the atmosphere at the location of interest.

This parameter and its constituent parts are very important in the problems under consideration. Given τ_o and some knowledge about the distribution of large scatterers in the local atmosphere, we can satisfactorily express the attenuation along most of all other paths in the local atmosphere. Additional information is needed to determine the noise due to sunlight, but here, too, τ_o and its constituent parts are of basic importance.

Closely related to the parameter $\tau_o = \tau(\ell_o)$ is the function $\tau(\ell)$, the optical thickness to altitude ℓ; the calculation of this function will be the major endeavor of this paragraph. The function $\tau(\ell)$ can be readily expressed in terms of the more common function $\beta(\ell)$, the attenuation per unit length at altitude ℓ integrating the well-known attenuation equation

$$\frac{dI}{d\ell} = -\beta(\ell)\, I \qquad (1)$$

along a path from above the atmosphere to ℓ, we get

$$I = I_o \exp\left\{ - \int_{\ell}^{\infty} \beta(\ell)\, d\ell \right\} \qquad (2)$$

so that

$$\tau(\ell) = \int_{\ell}^{\infty} \beta(\ell)\, d\ell; \quad \frac{d\tau}{d\ell} = -\beta(\ell) \ . \qquad (3)$$

Now, β is the scattering coefficient per unit thickness of air. Since there is negligible interaction between particles in air, we may represent β as the sum of scattering coefficients for the various constituents of the air. But by (3), we may also consider $\tau(\ell)$ as a sum of constituent parts. Such a breakdown will be very useful in the following discussion.

For ℓ less than a few kilometers, $\tau(\ell)$ can vary over orders of magnitude at the same location, depending on meteorological conditions; it may also vary markedly with frequency. For example, in the visible range, $\tau(0)$ ranges from a few hundredths to the orders of tens or higher. There are basic correlations between $\tau(\ell)$ for given ℓ and the geographical location and type of air mass present: $\tau(\ell)$ tends to be greater for higher ground level, decreases with distance from man-made centers of pollution, and tends to be greater in the presence of polar air masses than in the presence of tropical masses.

We note that $\tau(\ell)$ is a logarithmic measure of loss. To convert this to decibel loss, we simply multiply by $10 \log_{10} e = 4.34$.

7.1.5 Attenuation Due to Rayleigh Scattering

The Rayleigh optical thickness τ_R is a function of altitude above sea level; it is fixed except for seasonal and local variations too small to be of concern here. It was studied by Deirmendjian[14], with additional calculations in the C-D-S tables[13], and we shall summarize the results here. The Rayleigh optical thickness to sea level, $\tau_R(0)$ is plotted in Fig. 7.3 as a function of wavelength; in the visible range, it varies from about 1.0 in the violet to about 0.03 in the red.

The numberical tabulations of the cited authors can be represented in good approximation by the following empirical formula:

$$\tau_R(0) = 8.5 \times 10^{-3} \lambda^{-4.08} \qquad (4)$$

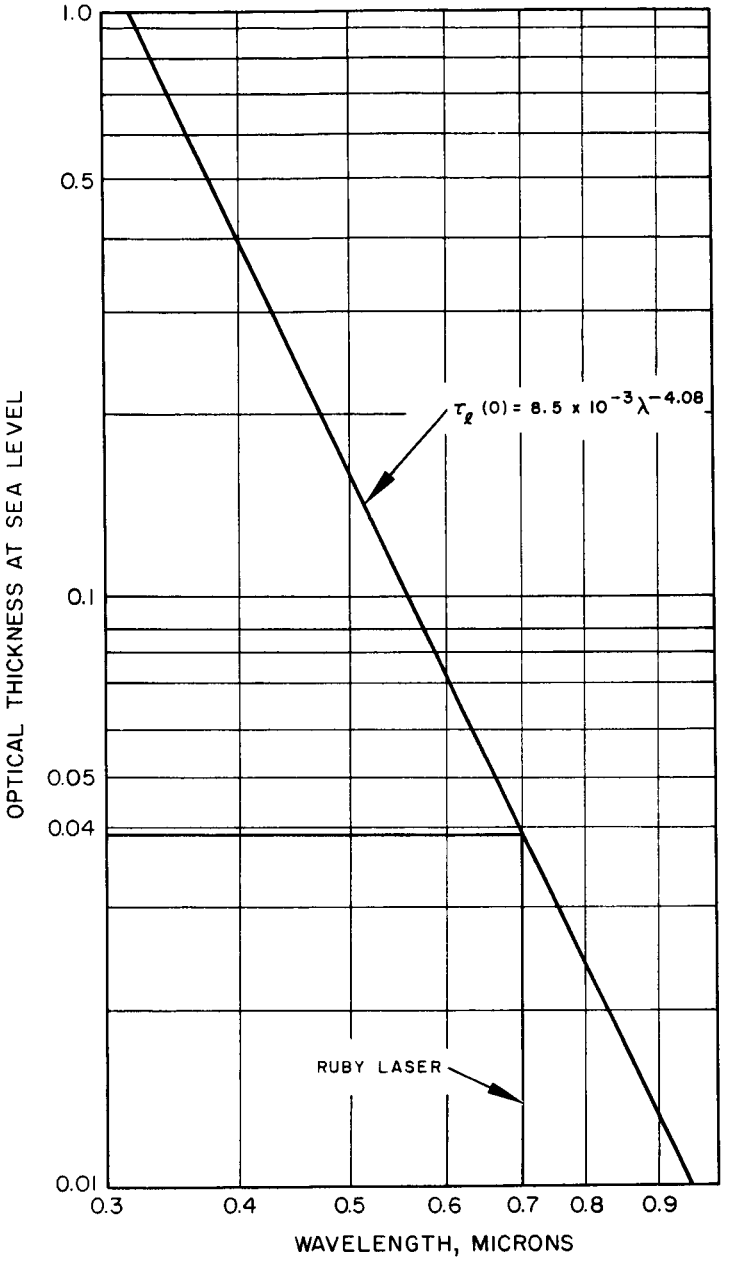

Figure 7-3. Optical Thickness of Sea Level due to Rayleigh Scattering

Here λ is wavelength in microns (this unit is convenient for calculations), and ℓ is measured in kilometers.

7.1.6 A Note on Absorption

When absorption is present, it can be accounted for as another additive component τ_A of τ. In the light and near infrared regions of the spectrum most absorption is due to H_2O vapor, CO_2 and O_3, although other gases have minor effects. In the visible range the absorption is negligible; what little is there is caused mainly by O_3 in the ionosphere. In the near infrared, the effect of O_3 is important only between 9 microns and 10 microns. Water vapor causes extremely high absorption in quite a few bands in the infrared, and CO_2 contributes a band around 4.5 microns.

For infrared propagation, the wisest thing is to avoid the heavy absorption bands and use the "window" frequencies where absorption is minor. These windows are given in any standard work on infrared, or in Chapter 16 of the Handbook of Geophysics[1]. A typical plot is given in Fig. 7.4.

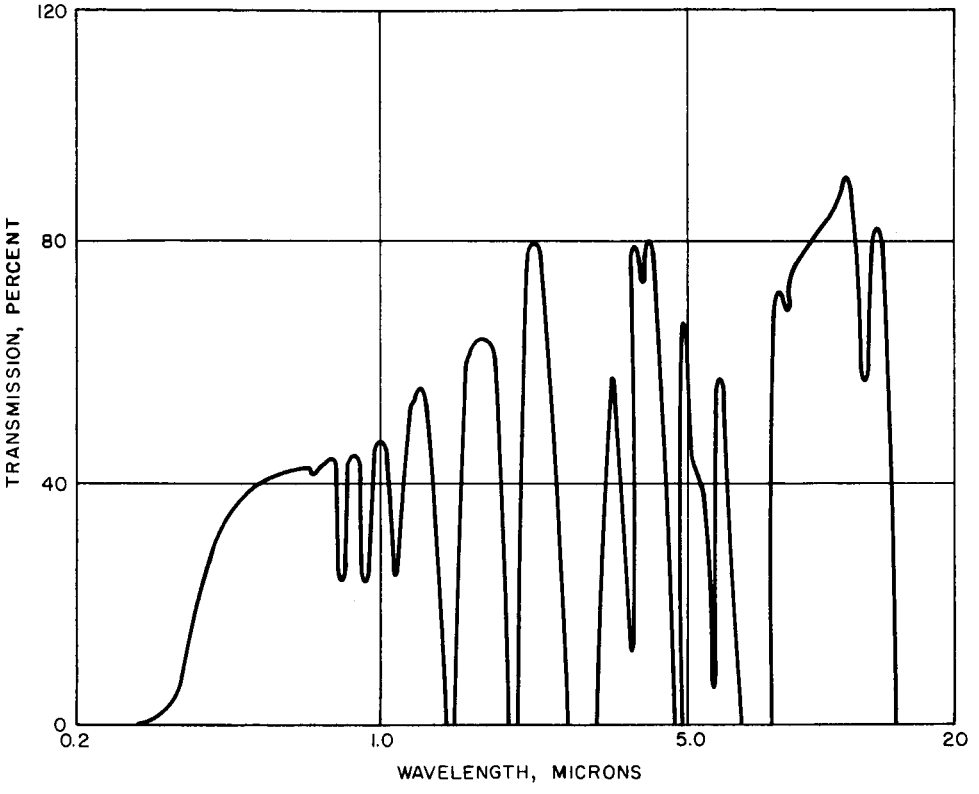

Figure 7-4. Atmospheric Transmission

7.2 THE TRACKING AND POINTING PROBLEM

A necessary condition for any communication problem is the establishment of an energy propagation path between the source and destination points.

Thus, any attempt to communicate with a space vehicle using directional antennas is underlined by the fact that the antennas should be within the beamwidth of each other. The problem becomes more and more difficult as the beamwidth (field of view) of the antennas becomes smaller and smaller. On the other hand it is clear that the smaller the beamwidth of the antennas is, the less the total power required to be radiated at the transmitting end in order to assume a certain power level at the receiving end. This high directivity is one of the attractive features which the laser presents as a communication instrument. However, the utilization of the extremely narrow beamwidths possible presents the serious problem of aligning the antennas. A clarification of this problem and its relation to the over-all system performance is given in this section.

7.2.1 Assumptions

The following assumptions have been made:

A. The position of the deep space vehicle (DSV) is known at long distances to an angular uncertainty γ smaller than the half beamwidth θ_b. This assumption is reasonable in view of the experience of DSIF (Deep Space Instrumentation Facility) operations, and the laser beamwidth under consideration.

B. The beamwidth $2\,\theta_b$ is the angle between half power points in the radiation pattern of a diffraction limited circular aperture, i.e.

$$\theta_b = 1.02\,\frac{\lambda}{d} \qquad (1)$$

Where, λ is the operational free space wavelength, and d is the diameter of the antenna under consideration. Moreover the radiation power is assumed to be uniformly distributed over the beamwidth. Observe that the same assumptions apply for the receiving antenna (dual of the transmitting antenna).

This assumption is justified due to the observed extremely small bandwidth and spectral purity of the laser radiation.

C. The pointing uncertainty θ_p of the antenna is the amplitude of the angular displacement over which control cannot be exercised. Hence it is assumed that the position of the antenna is equally likely to be at any point within $\pm \theta_p$ from the nominal (ordered) value.

This assumption is artificial and equally well we may assume any symmetrical distribution about the nominal position of the antenna. Here, however, we assume the simplest in order to stress the principles involved rather than computational techniques.

7.2.2 Discussion

Let us now examine the implications of the assumptions made. Observe that by assumption (A), if it was possible to position the antenna without error, the DSV could be acquired with probability one. This is because the beam illuminates a cone of half angle θ_{b_1}, larger than the angular uncertainty γ of the DSV position.

However the uncertainty of pointing the antenna accurately gives a probability that the DSV will be illuminated which is less than one if $\theta_{p_1} > 2\ \theta_{b_1}$.

This probability is the product of the probability that the DSV is in within the nominal angle $\pm \theta_{p_1}$ (which is one because $\theta_{p_1} > \gamma$) times the probability that it is within the beam, which is

$$P_1 = \frac{\theta_{b_1}}{2\ \theta_{p_1}} .$$

Now assuming that the DSV is illuminated, it is possible to find the direction of propagation of the radiation to a degree of accuracy, δ . Thus the on-board antenna may be pointed in this direction with a pointing accuracy $\theta_{p_2} > \delta$. Hence the probability of aligning the DSV antenna to the ground antenna is:

$$P_2 = \frac{\theta_{b_2}}{2\ \theta_{p_2}}$$

where again we assume that the target (here the ground antenna) is within the cone of angular uncertainty of pointing the DSV antenna, i.e. $\theta_{p_2} > \delta$.

Therefore the probability that the two antennas are within the beam of each other is

$$P = P_1 \cdot P_2 = \frac{\theta_{b_1} \cdot \theta_{b_2}}{4\,\theta_{p_1} \cdot \theta_{p_2}}$$

7.2.3 Radiated Power vs. Pointing Accuracy

The power P_r received by an antenna aligned to a transmitting antenna is given by the well known relation

$$P_r = P_t \frac{A_r}{\pi(\theta_{b_1} R)^2} \tag{2}$$

where P_t = transmitted power

θ_{b_1} = transmitted beam cone half angle (diffraction limited see e.g. (1))

A_r = receiving antenna aperture

R = distance between the two antennas.

Notice that assuming a diffraction limited antenna on the receiving end and a beam cone half angle θ_{b_2}, we have

$$\theta_{b_2} = 1.02 \frac{\lambda}{d_2} \tag{3}$$

where d_2 is the diameter of the antenna receiving aperture.

Then
$$A_r = \frac{\pi d_2^2}{4} = \frac{\pi}{4} \left(1.02 \frac{\lambda}{\theta_{b_2}}\right)^2 \tag{4}$$

Substitute in (2) solved for P_t, to obtain:

$$P_t = P_r \frac{4\,(\theta_{b_1}\,\theta_{b_2} R)^2}{1.04\,\lambda^2}$$

Now the minimum required power level may be given as a product, called signal to noise ratio (SNR), of the noise power which here, assuming that background and detector noise is negligible compared with signal (photon) noise, is given by

$$N = h \frac{c}{\lambda} B$$

where: $h = 6.6252 \times 10^{-34}$ watts, is the Planck's constant

$c = 3 \times 10^{8}$ m velocity of electromagnetic energy propagation in vacuum

and

B = operational bandwidth

Therefore

$$Pr_{min} = (SNR)_{min} N = (SNR)_{min} h \frac{c}{\lambda} B$$

Hence

$$P_{t_{min}} = (SNR)_{min} \frac{4 h c B}{1.04} \cdot \frac{(\theta_{b_1} \theta_{b_2} R)^2}{\lambda^3} \qquad (5)$$

Here we may assume that we have the same antennas at both ends, i.e.

$\theta_{b_1} = \theta_{b_2} = \theta_b$ to obtain:

$$P_{t_{min}} = (SNR)_{min} \frac{4 h c B R^2}{1.04} \cdot \frac{\theta_b^4}{\lambda^3} \qquad (6)$$

Now it is clear that the smaller the beamwidth, the smaller the power required to be transmitted, the rest of the parameters being held constant. However, this cannot be done indefinitely, the main obstacle being the pointing accuracy.

From the previous discussion is evident that the beam cone half angle θ_b should be larger than the pointing uncertainty, θ_p.

That is, $\theta_b \geq K \theta_p$.

The value of K depends on the particular distribution of the random displacement of the antenna from its nominal value and the desired probability to illuminate the target successfully.

Hence we get from (6)

$$P_{t,om} = (SNR)_{min} \frac{4 K^4 h c B R^2}{1.04} \frac{\theta_p^4}{\lambda^3} \qquad (7)$$

This equation makes clear the importance of improving the pointing accuracy, for the saving in transmitting power improves as the 4th power of any improvement made on θ_p. Fig. 7.5 illustrates this relationship.

Caution should be exercised in interpreting equation (7) in terms of λ. It appears that it is advantageous to go to lower frequencies (longer wavelengths). However, lower frequencies mean larger apertures for fixed gain and thus greater weight. Moreover, detector sensitivity, for existing detectors, falls off even more rapidly than the cube with wavelength.

It seems that optimum wavelength may be in the near IR region. In relation to pointing accuracy the present state of the art leaves much to be desired. For economincal application of lasers, and deep space communication links in general, improvements of the pointing accuracy and attitude control of order of magnitude or more of the present state (assume to be 10 μ rad) are required. For such extreme accuracies in sensing and controlling, the laser provides a promising instrument. Interferometer techniques could perhaps be applied, in connection with the coherent, highly monochromatic radiation available, to measure angles and angle rates with increased precision.

Also array phasing techniques may be employed to position accurately a laser beam, provided attitude control of the same order of magnitude can be made available (It is of interest here to note that the OAD requires attitude control to 0.2 seconds of arc).

Finally, improved detector sensitivity at the longer wavelengths would be quite profitable if such can be obtained. This desirability stems essentially from the inverse cube dependence on wavelength which results from the increasing photon noise at higher frequencies.

REFERENCES

1. U.S.A.F., Handbook of Geophysics. Geophysics Research Directorate, Air Force Cambridge Research Center, 1957, First Edition.

2. Elder and Strong, "The Infrared Transmission of Atmospheric Windows", Franklin Institute, Journal, 255 (189) March 1953.

3. Gebbie, et al, "Atmospheric Transmission in the 1 to 14 Micron Region," Proc. Roy Soc. (London), A206, P 87, 1951.

4. Gibbons, "Wavelength Dependence of the Scattering Coefficient for Infrared Radiation in Natural Haze," Jour. Opt. Soc. Amer., 48(172), March 1958.

5. Howard, Burch & Williams, "Infrared Transmission of Synthetic Atmospheres," Jour. Opt. Soc. Amer., 48(186, 236, 242, 334, 452), 1956.

6. Knestrick et al, "Atmospheric Scattering Coefficients in the Visible and Infrared" Jour. Opt. Soc. Amer., 52 (1010) Sept. 1962.

7. Larmore, "Transmission of Infrared Radiation Through The Atmosphere," Office of Naval Research, Proceedings of the Infrared Information Symposiums. 1 June 1956, 2 Dec. 1956.

8. Penndorf, "Tables of the Refractive Index for Standard Air and the Rayleigh Scattering Coefficient for the Spectral Region between 0.2 and 20.0 Microns and their Application to Atmospheric Optics," Jour. Opt. Soc. Amer., 47 (176) Feb. 1957.

9. Summer, Photo Sensitors, Chapman and Hall

10. Middleton, Vision Through the Atmosphere, University of Toronto Press, 1952.

11. Taylor and Yates, "Atmospheric Transmission in the Infrared," Jour. Opt. Soc. Amer., 47 (223), 1957.

12. Yates, "Total Transmission of the Atmosphere in the Near Infrared," NRL Report 3858, Sept. 10, 1951.

13. Coulson, Dave, and Sekera, Tables Related to Radiation Emerging from a Planetary Atmosphere with Rayleigh Scattering, (Univ. of Calif. Press, Berkeley, 1960.

14. D. Deirmendjian, "The Optical Thickness of the Molecular Atmosphere," Archiv. Meteor., Geophys. Biok, Series B, 6, 452 (1955).

15. Lengyel, B.A. and Mitzner, K.M., "Some Effects of the Atmosphere on a Laser Radar System," Hughes Research Report #217, Oct. 1961.

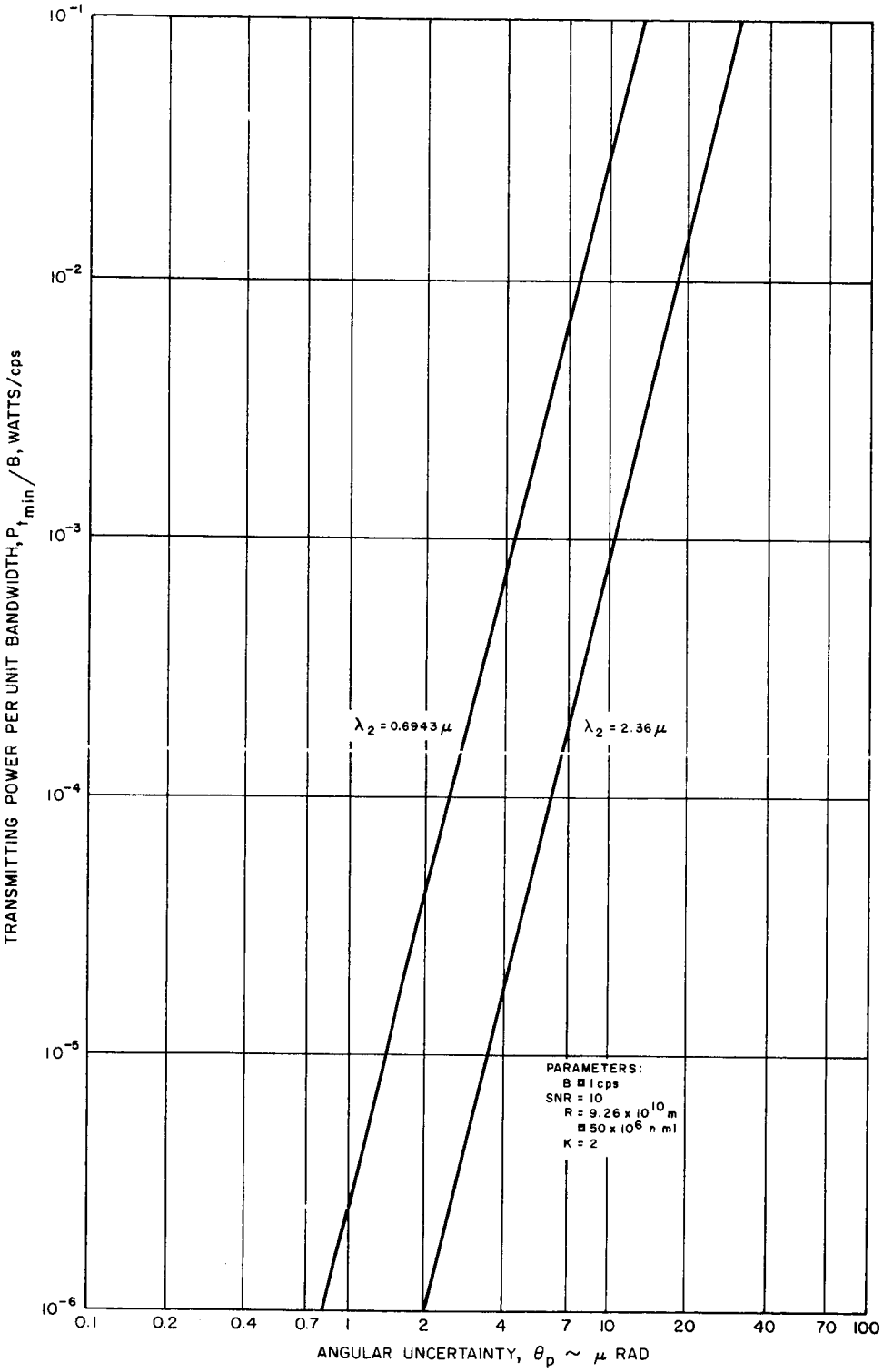

Figure 7-5. Effect of Pointing Inaccuracy

8.0 SYSTEM EVALUATION

The following discussion delineates some of the major factors that contribute to the selection of the best system from communication links considered in this study.

Each communication link is divided into two parts by direction of transmission. The transmitter power required at each station to overcome the background noise is also given. In some instances it is not the limiting factor while in others the power needed exceeds the maximum power capabilities of the present state-of-the-art, thereby making a particular link less desirable by forcing a reduction in the capability by narrowing the beamwidth or limiting the operating time.

Table 8.3 gives the values for the sensitive parameters but in general, the obvious conclusions derived from the table are insensitive to the state-of-the-art. As an example, consider a breakthrough in the maximum power capabilities of lasers. This would apply equally well to all of the links and it is not likely to change any decision as to the best system to develop.

The next phase of this study will be a continuation of this evaluation phase.

8.1 TRANSMISSION LINKS

8.1.1 From Moon to Deep Space Vehicle

The ecliptic missions put the moon in an advantageous position for a relay station. It orbits the earth about once every 28 days at an inclination of about 5^{o} to the equator which in turn is inclined with

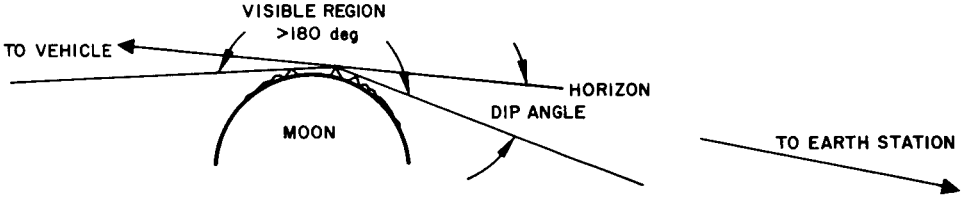

Figure 8-1. Lunar Pole Field of View

respect to the ecliptic about 23-1/2°. The moon's axis of orientation and rotation rate are such that it keeps the same face toward the earth. The combination of inclined orbit and axis orientation makes a couple of regions on the lunar surface visible from both the earth and most of the deep space vehicle positions in the ecliptic. These are at the lunar poles. Allowance can be made for the fact that the moon's polar axis is not exactly orientated perpendicular to its orbit plane when selecting a particular spot in the vicinity of the pole. If a high mountain can be selected at the pole that is isolated from other mountains equally high by a few miles, considerably more than a hemisphere of visibility could be made available, (Figure 8.1). This is principally due to the diameter of the moon being small allowing a large depression angle (DIP) of equal altitude points when only slightly removed in distance. By taking advantage of this relationship only one moon station is needed to continuously see both the earth and the deep space vehicle.

Many advantages of a moon base are immediately apparent. Continuous coverage of deep space operations are possible. All transmittal data from the vehicle may be sent directly to the ground without necessitating any bulky storage capacity at the source. The moon base, not being limited by weight (multiple vehicles can deliver the required equipment) can if necessary process or store the information for future use.

For a minimum equipment of moon and space vehicle link, more than one earth station is needed to provide continuous coverage. Three are satisfactory to provide a real time TV reception such as would be required for a hard intercept (impact) of Mars by an instrumented probe. For only one ground station and no loss of instrumentation data in the link, about 13 hours of storage data are needed.

Only one mission plane, the selection of which is very unlikely, will not provide continuous coverage capability. This is the one involving the earth - moon orbit. Here once each month the moon will be occulted by the earth, but only for a brief 21 minutes. This is only 0.05 percent of the total time, hardly worth considering.

There is a small probability of other planets occulting the moon for extremely long range missions. For example, on a mission to Jupiter, Mars would block the view for only a few minutes every 687 days and then only if the two intersecting or coinciding planes contained all three bodies (Moon, Jupiter and Mars).

The other possibility is the interference from the sun. This interference (if restricted to occultation only) would occur only at opposition (moon and deep space vehicle on opposite sides of the sun).

The full blanking would occur only during the time the sun's 1/2 degree disk coincides (assume 1 astronomical unit distance to sun), but the sun's appreciable atmosphere (corona) extends out a considerable distance making the apparent sun's disk much larger. A mission of the type that would give this condition would probably be one where apparent change in relative position of the moon with respect to the sun would be quite slow requiring several weeks without communication. This type of mission is not likely to be used.

There may be a mission where one lunar pole is favorable for some part of the time and the other pole more favorable for the remainder of the mission. This could require two relays set up, one on each pole but only one used at a time.

Background

Several types of background noise exist depending on the location of the two stations. The deep space vehicle, when looking at the illuminated side of the moon for a transmitted signal, must discriminate the signal from all the reflected sunlight, this is the worst case i. e., the vehicle is in such a position that its entire field of view is illuminated. Assuming a beam width of 10 microradians, the field of view will include a circular section of the moon's surface 500 miles in diameter. The illuminance of the moon when viewed at earth distance away (3.84×10^8 meters) is 0.292 lumens per square meter. (LUX). Extending this out to 50 million miles the moon appears as 7.5×10^{-13} watts per square centimeter (1 lumen ▫ 1.496×10^{-3} watts). This says that for a signal to noise ratio of 10, the lunar transmitter must send in the direction of the space vehicle sufficient energy to impart on a surface 50 million miles away 7.5×10^{-12} watts per square centimeter. Assuming the moon transmitter has an antenna gain of 115 db as given in table 1.3 of the previous report, and the receiver uses a 10 angstrom bandwidth filter removing 99.9 percent of the noise power at the frequency selected, the transmitter will have to send out 25.6 watts. If we assume a reasonable power maximum of only 10 watts the signal to noise ratio or bandwidth is reduced proportionally.

Actually, this is very pessimistic. The moon is not always full but goes through phases depending on the location of the space vehicle and sun. The amount of background illuminance falls off rapidly as the phase angle increases. Figure 8.2 shows the amount of power required to overcome the background noise. If the power limit is 10 watts, only slightly more than 21 percent of the time will the signal to noise ratio (S/N) be less than 10. For most outside earth orbit missions the S/N will never be less than 10. Besides the favorable diminishing phase angle, the entire field of view will not always be covered by reflected sunlight. For missions in the ecliptic, somewhat less than half of the field of view will be illuminated due to the preferred

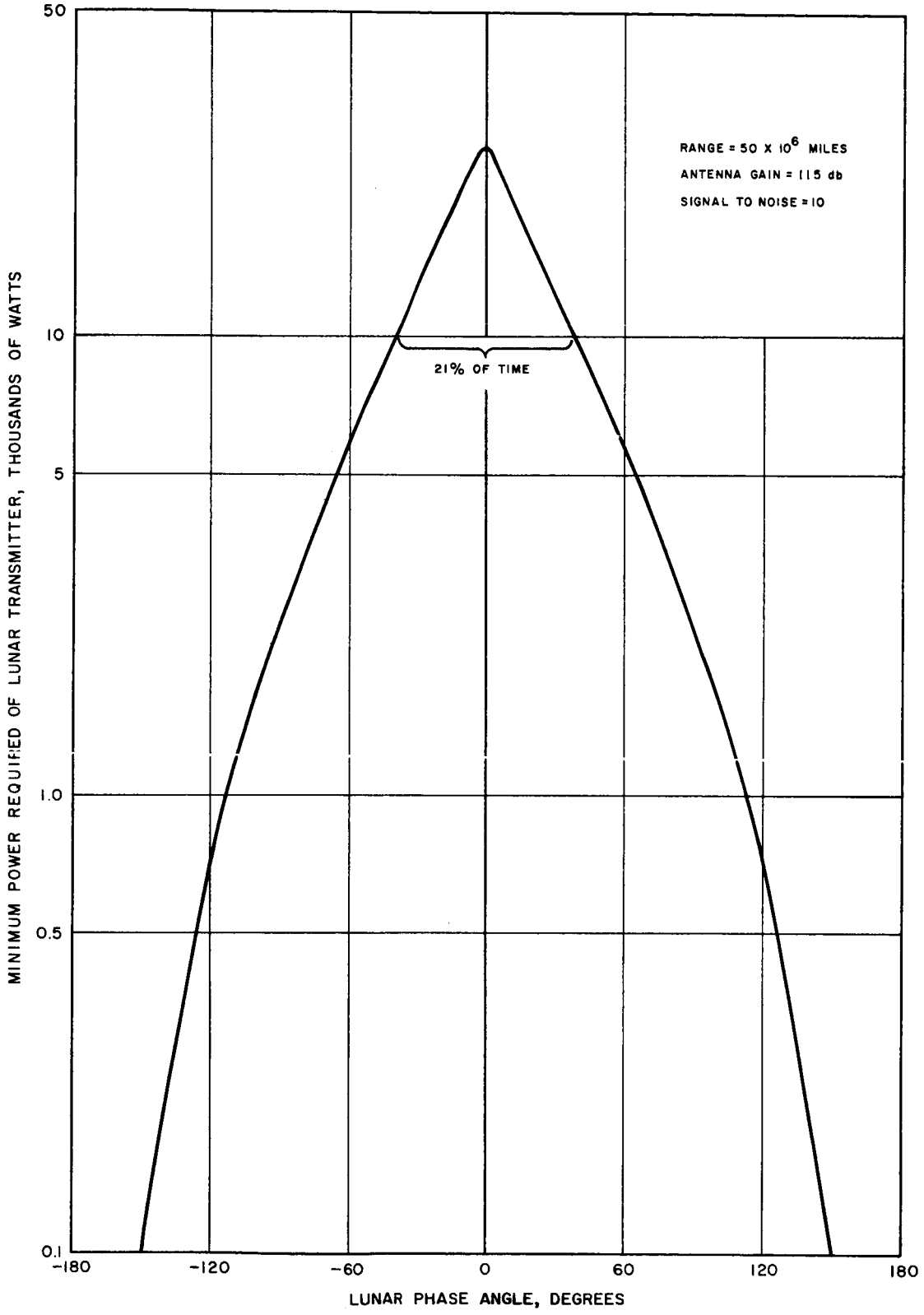

Figure 8-2. Power Required to Overcome Average Lunar Background as a Function of Phase Angle

installation on the moon being at the poles where slightly more than half of the cone angle of acceptance misses the moon entirely.

The 10 angstrom bandwidth filter leaves a fraction of the background noise dependent on the frequency selected. Looking at the spectral distribution of the reflected sunlight (Figure 8. 3) a selection of a laser frequency less than 0. 4 microns or more than 0. 7 microns gives a factor better than 1 part per 1000. For example picking 2. 36 micron, the frequency of $Ca F_2 Dy^{2+}$, gives a noise irradiance of only 1 part in 20, 000. However, detector sensitivity diminishes more rapidly with λ than this effect, and for shorter λ, signal noise increases, adequate lasers are not yet available, and detector sensitivity does not improve significantly.

Because the moon station will be on the apparent edge of the moon, the background may at times contain the earth which has an albedo 4. 92 times as great as the moon. This means that for the same conditions above, the transmitted power required will be more than 76 watts, but only for 72 minutes of each month, and then only if the orbital plane of the moon-earth and moon-vehicle line coincide. It is possible that the sun will be in the background which would prohibit communication. This is not likely to occur during most missions.

The lunar station looking at the space vehicle for a signal does not have to contend with the high background most of the time. This lowers the vehicle transmitter power requirement and (or) antenna aperture size. Some missions may place one of the other solar system bodies in the background for part of the time. A mission to Mars under the worst background conditions, "full" Mars in view, will require a transmitted power of 22. 6 watts with the same assumptions of antenna gain etc. To overcome the background noise of Jupiter only 0. 134 watts of vehicle transmitted power is required. If inferior missions (inside the earth orbit) are selected, such as to Venus the sun will only illuminate the other side of the planet and the moon based station will look at the dark side of the planet for a background, and will not see a "full" Venus except for the special mission at opposition.

8. 1. 2 From Deep Space Vehicle to Moon

The same arguments hold for transmission in this direction as in the reverse direction just mentioned. It is assumed that larger, more sensitive equipment can be mounted on the lunar surface than on the space vehicle. The space vehicle may not have the high power transmission capability, but the lunar based receiver can have a larger antenna capture area. A twelve inch mirror on the space vehicle can be used while a 40 inch is not unacceptable for the lunar mounted equipment. The one big difference is the amount of power needed to

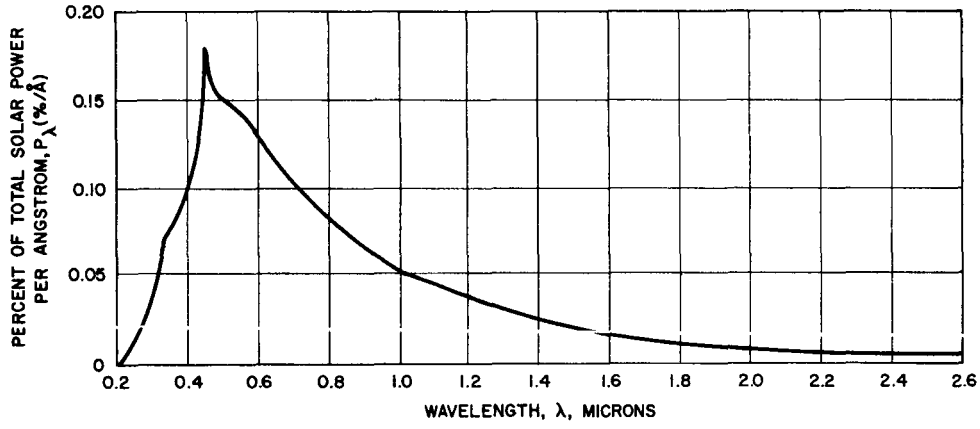

Figure 8-3. Solar Spectral Density

overcome the background noise. Most noise, when viewing the vehicle from the moon, will be from the star background. The greatest concentration of stars is in the direction of the galactic equator. There are about five times as many stars per unit solid angle in this direction as the average over the whole sky. In the direction of the galactic poles, the density of the stars is only about 1/5 the average. For missions in the ecliptic, where all out planets lie, the regions of high star density will be encountered about once every 6 months. Since the galactic plane has considerable width, a 6 month mission will have the galaxy in the background for about a month. Any communication link to deep space must consider the worst case in determining the transmitted power required to overcome this background.

The noise energy distribution is not uniform, but consists of strong discrete sources with almost nothing in between. With a small field of view the probability of an interfering star depends on the density distribution of the stars which have sufficient noise power to serve as sources of signal interference. Figure 8.4 shows a plot of the stellar illuminance as a function of the number of stars per square degree. From this curve it can be seen that the probability of an interfering encounter increases with increased sensitivity. By setting the probability of interference at some acceptable value like $P = 0.01$ (i.e., 1% of the time noise will be a problem) and picking the field of view of 10 microradians, the noise power is found to be 2.5×10^{-20} watts per square centimeter from the direction of the galactic center setting the signal to noise ratio at 10 gives a received signal power requirement of 2.5×10^{-19} watts per square centimeter. This is the amount of power per unit area that must arrive at the receiver from a transmitter located 50 million miles away. The amount of power transmitted is determined by the equation:

$$P_T = \frac{4 \ R^2 P_R}{G_T \ K} \qquad (1)$$

where

P_T = power transmitted, in watts

P_R = power received, in watts/cm^2 (background noise times S/N)

R = range, in cm

G_T = gain of transmitting antenna (116 db)

K = filter constant (1000 for 0.6943 microns)

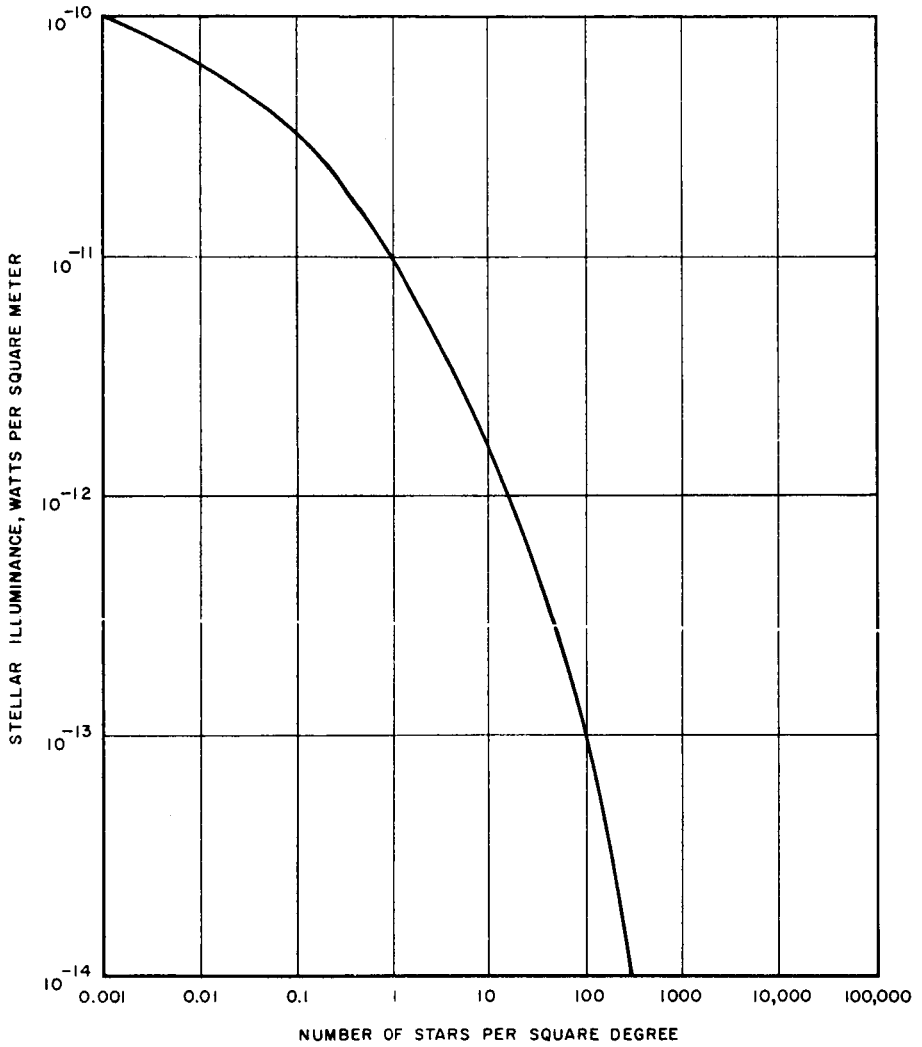

Figure 8-4. Stellar Power Density

The gain of the transmitting antenna is determined from its size and the frequency used. It is assumed that the vehicle antenna will be restricted to a diameter of 12 inches and the frequency is 0.6943 microns giving a gain of 116 db or 3.16×10^{11}.

The spectral distribution of the stars is assumed to be the same as for our sun (shown in Figure 8.3). A filter is used to remove the energy outside the desired band. For 0.6943 microns, the filter bandwidth was assumed to be 10 Å, corresponding at that frequency to a filter constant K of 0.001 part of the total stellar radiation.

The required transmitter power is then found to be 8.58×10^{-7} watts. Since the expected transmitter power available is many orders of magnitude greater, the probability of encountering star background interference becomes negligibility small.

Other types of noise also should be considered. In addition to the possibility of occulting from the earth which is identical to that mentioned in section 8.1.1 there is the chance that a planet will be in the background. This is a very remote possibility and if it did occur it would only last for a few minutes at most.

8.1.3 From Synchronous Satellite to Deep Space Vehicle

This transmission beam, like the one in reverse, is obstructed by the moon and earth occulting the space vehicle from view, but only when the space vehicle is in the plane of the satellite and respective body. The moon can only interfere once a month and then only for about one hour at worst, whereas the earth may interfere once a day for about 1 - 1/4 hours. These are easily avoided by selecting the right inclination angle for the satellite thereby allowing continuous coverage for the entire lifetime of the average 6 months mission.

The background noise, when looking at the satellite from deep space is principally dependent upon the star density. The worse case for star background is when the center of the milky way (our galaxy) is directly behind the satellite. Assuming the same equipment is used on the satellite as it is on the space vehicle in terms of antenna gain and frequency, the satellite must transmit the same power as the deep space vehicle does, i.e. 8.58×10^{-7} watts, to give S/N = 10 against the background noise.

In addition to the occulation of the satellite by the earth each day for 72 minutes,(for the special mission when the deep space vehicle is in the orbit plane of the satellite), an equal time will be spent in transit with the earth as a background. When the earth is full (the worst case), the power required for a signal to noise ratio of 10 goes up to 130.5 watts. The earth goes through phases where the

average background power diminishes. Figure 8.5 shows a plot of earth phase angle as a function of power required to overcome it as a background. It can be seen from the curve that if the maximum transmitted power is limited to 10 watts, the signal to noise ratio of 10 or more will occur for 50% of the time of transit. Another way of saying this is that reception is possible when the space vehicle-satellite-sun angle is greater than $\pm$ 90 degrees. The only other large background source is the sun; communication against this source is impractical due to the extremely large power requirements. It may be possible in the future to select a frequency either much higher or much lower that will permit the sun to be in the background and still permit moderate transmission power levels. Until then the mission profiles must be selected to avoid the sun.

8.1.4 From Deep Space Vehicle to Synchronous Satellite

The selection of a satellite for a relay has the advantage of not being blocked from the view of the deep space vehicle very often. If the right orbital inclination is selected, the satellite can see the space vehicle continuously for several months. For low altitude satellites this inclination angle approaches the normal to the earth-space vehicle line of sight. When the orbit precesses or the space vehicle moves to a position in the orbital plane, the space vehicle is occulted for nearly half the time. If the orbital altitude is increased the percent of time occulted is reduced as well as the inclination angle required for continuous coverage. This increased height also reduces the number of ground stations needed for continuous coverage. When synchronous height (22,300 miles) is reached, only one ground station is needed. If an equatorial inclination is selected (23-1/2 degrees with respect to the ecliptic) most missions in the plane of the ecliptic, where all the planets lie, will have continuous coverage for about 6 months. Those two times a year, when the deep space vehicle will be in the orbital plane of the satellite, occultation will occur due to the earth once each day for a maximum of 72 minutes. The moon will at times interfere with communication in a similar manner but never more than once a month and then for less than an hour.

The background noise will be a maximum when the galactic center is behind the deep space vehicle. The noise power from the galactic center is found to be 2.5×10^{-20} watts per square centimeter (section 8.1.2). The signal power needed to overcome this noise with a signal to noise ratio of 10 is 2.5×10^{-19} watts per square centimeter. Computing the amount of transmitted power required of the space vehicle from equation (1) gives 8.58×10^{-7} watts.

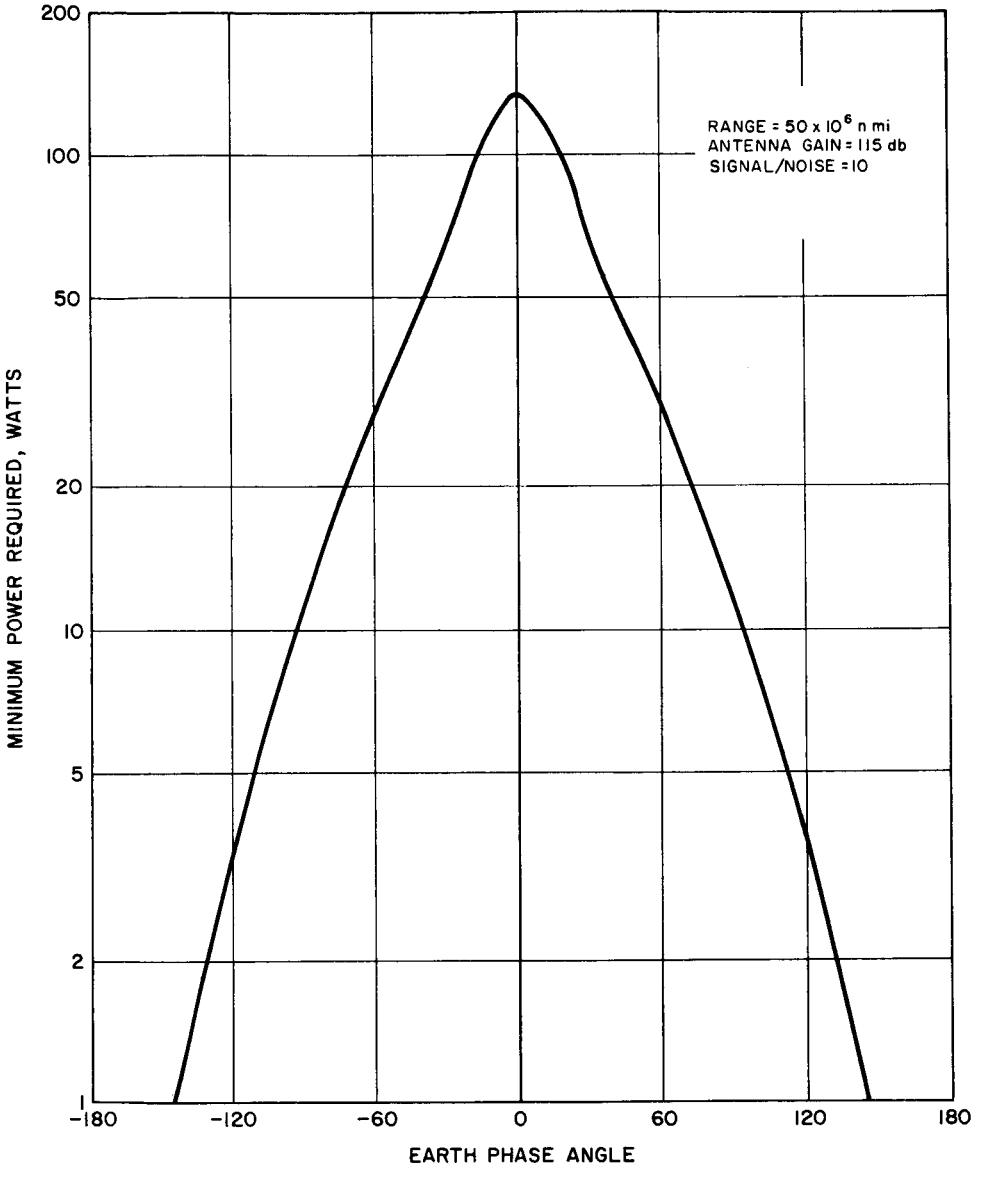

Figure 8-5. Power Required to Overcome Earth Background as a Function of Phase Angle

8.1.5 From Earth to Deep Space Vehicle

The atmospheric effects play a big role in this path length. They are discussed in part in section 7 of this report and in section 2.3 of the previous progress report.

A single ground station will be able to transmit to the space vehicle for less than half time due to the earth rotation moving the station out of view. If the station location on earth is carefully selected, clouds should not appear as a severe interference problem. The space vehicle may periodically be eclipsed by the moon, but the probability of this occurance is quite low.

Background Noise

The satellite must look at an average earth background when 'full' of 3.69×10^{-12} watts per square centimeter. This is based on an average albedo of 0.29 which is true for bare desert areas. Actually the albedo varies considerably depending on the region viewed. It may be as low as 0.04 for recently turned rich black soil, or as high as 0.76 for glazed snow. Table 8.1 shows the albedo for different degrees of ground surface. The desirable conditions for good receiving involve long periods with no clouds which accurately describe most of the desert regions. The clouds themselves, when present, offer a high apparent brightness as shown in Table 8.2.

Using the above background noise and assuming a desired signal to noise ratio of 10, the received signal power required is found to be 3.69×10^{-11} watts per square centimeter at the deep space vehicle. For this amount of power received, the power required at the edge of atmosphere is 123.2 watts with a filter factor of 1000/1. The daytime atmospheric image motion is involved here so the 10 microradian beam width should be replaced with something more realistic like a 50 micron radian one, giving an increase in power required of $25\ (50/10)^2$ to 2080 watts. Using an atmospheric attenuation constant of 0.5, the transmitted power then becomes 6160 watts. At night this is reduced to 1.17×10^{-7} watts for the 10 microradian beamwidth.

8.1.6 Deep Space Vehicle to Earth

Transmission directly to a single station on earth from the deep space vehicle can be accomplished for considerably less than half of the time primarily due to the earth turning on its axis and removing the station from view. The atmosphere makes reception on the ground near the horizon uncertain. The reasons for this uncertainty are given in the previous progress report in section 2.3 called "Communication Channel Characteristics". For continuous coverage, at

TABLE 8.1 Spectral Reflectivity of Solid Ground Surfaces

(After Krinov as Quoted in Handbook of Geophysics)

	WAVELENGTH
Type of Measurement and Type of Surface	6943
Measured from Ground:	
Fresh Fallen Snow	0.71
Snow Covered with a Film of Ice	0.76
Limestone, Clay, Similar Bright Objects	0.72
Sand, Bare Areas in Desert, Some Mountain Outcrops	0.29
Podzol, Clay, Loam, and Other Soils; Paved Roads and Some Buildings	0.17
Black Earth, Sand Loam, Earth Roads	0.04
Coniferous Forests in Winter	0.06
Coniferous Forests in Summer; Dry Meadows; Grass in General Excluding Lush Grass	0.14
Deciduous Forests in Summer, Lush Grass	0.18
Forests in Autumn, Ripe Field Crops	0.31

TABLE 8.2. Reflectivity of Various Cloud Types

(Fritz)

CLOUD TYPE	REFLECTIVITY
Stratocumulus, Overcast	0.56 - 0.81
Altostratus, Overcast	0.39 - 0.59
Cirrostratus and Altostratus	0.49 - 0.64
Cirrostratus, Overcast	0.44 - 0.50
Very dense clouds of extensive area and great depth	0.78
Dense clouds, quite opaque	0.55 - 0.62
Dense clouds, nearly opaque	0.44
Thin clouds	0.36 - 0.40
Stratus, Overcast, Thickness greater than 800 ft.	0.60 - 0.74
Stratus, Overcast, Thickness 600 to 800 ft.	0.38 - 0.70
Stratus, Overcast, Thickness less than 600 ft.	Usually 0.20 - 0.50 Occasionally higher or lower.

least 3 earth stations are required at about 120 degrees increments of longitude. These stations must be located away from areas of cloud cover or extreme atmospheric turbulance. In the United States, Arizona offers some locations that have 360 days a year clear weather. In Africa there are some areas where no moisture has fallen for a decade. Even multiple earth stations does not insure continuous coverage but nearly so. When the space vehicle is in the orbital plane of the moon there is a probability of 0.5 that the moon will block the space vehicle from view once each month for about 58 minutes. It is possible that other solar bodies could occult the space vehicle during very long range missions but the probability of this occuring is very low.

Background noise

The background seen by the earth station is quite varied. At low latitudes during the night it approaches the star background limit but during the day when the atmosphere scatters sunlight it is much higher. For a daytime cloudless sky using a 10 microradian beam width, a filter of 10 angstroms width, and the standard solar spectral distribution, the sky background looking up from the earth is 1.96×10^{-18} watts per square centimeter. If we desire a signal to noise ratio, of 10, we need power delivered at 1.96×10^{-17} watts per square centimeter at the ground. Assuming an atmospheric attenuation of 0.5 the power needed just above the atmosphere is 3.92×10^{-17} watts/cm^2. This assumes zero thickness atmosphere which is an acceptable approximation when considering a 50 million mile range. The transmitted power required is 0.134 watts.. A special problem appears with this type of link involving the atmosphere. The daytime atmosphere turbulance causes image motion sufficiently large to periodically result in complete loss of signal for a 10 microradian beamwidth. Since the power required is proportional to the square of the beam width, an increase from 10 microradians to 50 microradian causes a 25 fold $(50/10)^2$ increase in power. The 50 microradian beamwidth power requirement is then 3.34 watts.. If a beamwidth of 100 microradians is used, the power requirement is 13.4 watts.

The night sky background is about 5 orders of magnitude smaller giving a transmitted power requirement of about 4.25×10^{-7} watts for a 10 microradian beamwidth. If continuous coverage is not required or if multiple stations are permitted on the earth, the advantages of the low power requirements of the night sky can be used. Assuming transmission is possible starting 30 minutes after sunset and continues until the earth zenith look angle is limited to 75 degrees (15 degrees from horizontal), the useful phase angle can approach to within $\pm$ 157-1/2 degrees. These limits are due to twilight and bad seeing conditions near the horizon. Figure 8.6 shows the useful time

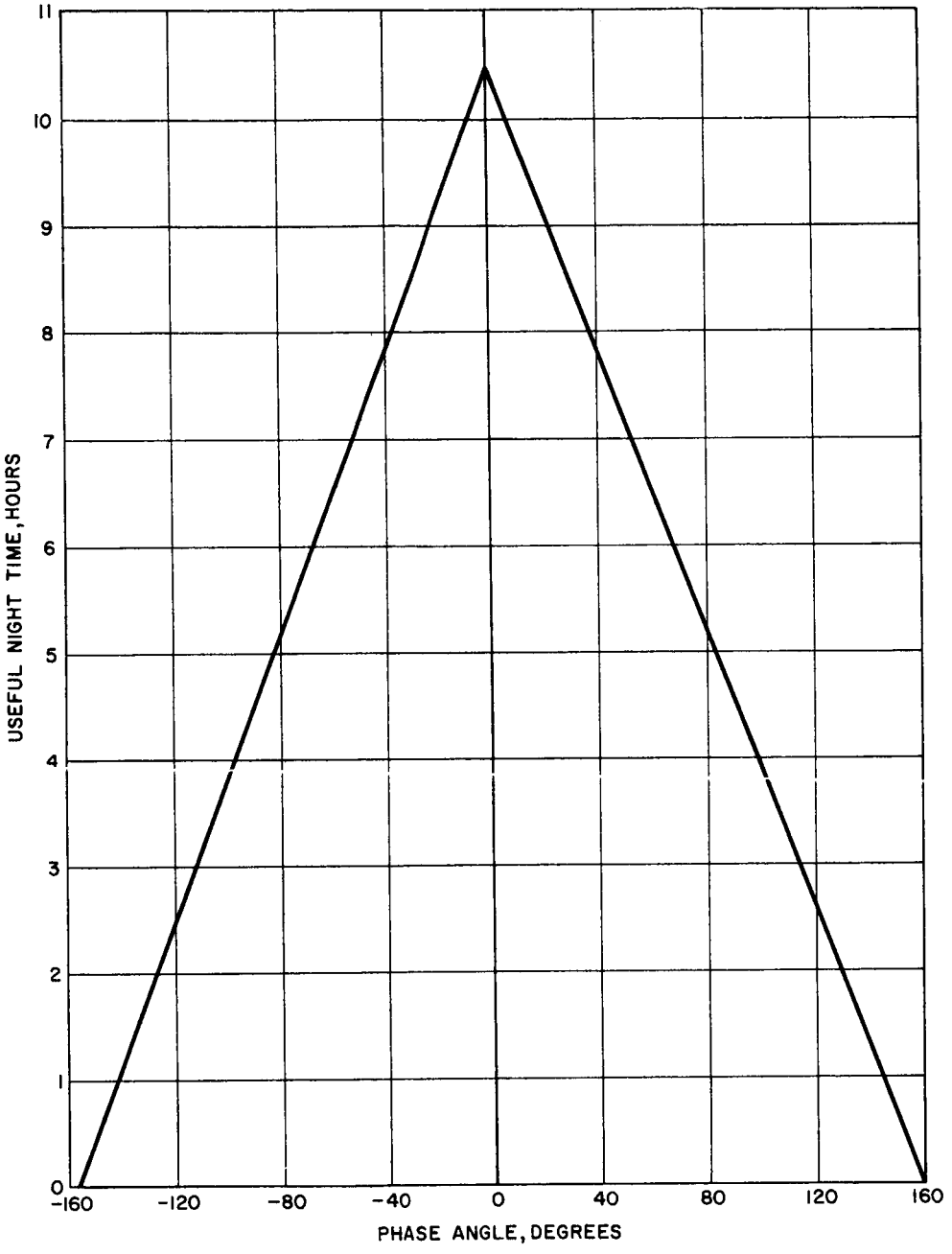

Figure 8-6. Useful Night Time of Single Earth Station

of a single earth station as a function of the phase angle. If continuous coverage is desired the number of equally spaced stations needed can be found by the ratio of 24/T where T is the time of a single station's usefulness in hours taken from Figure 8. 6.

TABLE 8.3. SYSTEM CHARACTERISTICS

LINK	Transmitter Power*	Background Noise (Total)
Moon to DSV	25.6 watts	7.5×10^{-13} watts/cm^2
DSV to Moon	8.58×10^{-7} watts	2.5×10^{-20} watts/cm^2
Sat. to DSV	8.58×10^{-7} watts	2.5×10^{-20} watts/cm^2
DSV to Sat.	8.58×10^{-7} watts	2.5×10^{-20} watts/cm^2
Earth to DSV	Day - 6160 watts (50 urad) Night - 1.71×10^{-6} watts (10 urad)	3.69×10^{-12} watts/cm^2
DSV to Earth	Day - 3.34 watts (50 urad) Night - 4.25×10^{-7} watts (10 urad)	1.96×10^{-18} watts/cm^2

* The power required to overcome the background noise with a signal to noise ratio of 10.

TABLE 8.4 SYSTEM COMPARISONS

LINK	ADVANTAGES	DISADVANTAGES
SATELLITE - Deep Space Vehicle	· No atmospheric effects (attenuation, refraction, image motion, beam spread, night-day sensitive. · Continuous coverage · Low background noise · Cost less than moon link	· No atmospheric protection · Occultated once a day when DSV in orbital plane · Limited antenna size · Cost more than earth link · Reacquisition more complex
MOON - Deep Space Vehicle	· No atmosphere effects (attenuation, refraction, image motion, beam spread, night-day sensitive) · Continuous coverage · Low background noise (one-way) · No reasonable antenna size limit · Reacquisition problem simplified	· No atmospheric protection · Occultated once a month when DSV in orbital plane · Most expensive · High background noise (full moon) · Additional ground stations
EARTH - Deep Space Vehicle	· Atmospheric protection · Logistics simplified	· Atmospheric effects (attenuation, refraction, image motion, beam spread, night-day sensitive) · Reacquisition more complex and frequent · Multi-earth stations for continuous coverage · High background noise (day)

8.2 CONCLUSIONS

Table 8,4 gives in tabular form the obvious advantages and disadvantages of the three communication links considered. What is not given is a list of subtle influencing factors that can change or make a system the favored one. As an example, judging the systems from the facts given, one could logically arrive at a preferred system involving a satellite to deep space vehicle link. If the mission is specified as being to Mars in the near future and continuous coverage is not needed, the body dynamics, cost, and reliability favors the earth-mounted station operating only at night. Had the mission selected been in the other direction, toward Venus for example, the power requirement alone to compete with the sunlit earth (no night sky) would have ruled out the possibility of a terrestrial based station.

If an earth to deep space vehicle system is selected, the best choice for operating frequency is made much more difficult. The atmospheric windows of lowest attenuation are not necessarily the best because some of these occur at very poor state-of-the-art detector sensitivity frequencies or in low power laser modes. An attempt to optimize the frequency selection for atmospheric and vacuum environment as well as the whole system will be a continuing effort during the following study phase.

The study of deep space environment and its effects on the electro-optical design will be completed and used in the final selection of the design.

The foregoing analysis makes it evident that improvement is required in various areas of system technology before efficient utilization of wide bandwidth optical communications systems can be made. Of principal concern is the problem of tracking and pointing of the extremely high gain antennas needed for communication over astronomical ranges. Moreover, increased detector sensitivity at longer wavelengths (IR) would serve to diminish the effect of the large signal (photon) noise which so seriously reduces bandwidth at the higher frequencies. In addition, resolution of the problem of achieving high signal quality CW laser operation at increased signal power must be solved. These problems will be addressed early in the design phase to establish an optimum system design which is commensurate with the growing state-of-the-art of the various components involved.